Calhoun: The NPS Institutional Archive
DSpace Repository

Theses and Dissertations 1. Thesis and Dissertation Collection, all items

1974-06

Analysis of some two-characteristic Markov-type manpower flow models with retraining applications.

Hayne, William J.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/16886

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

DUDLEY KNOX LIBRARY

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

ANALYSIS OF SOME TWO-CHARACTERISTIC MARKOV-TYPE MANPOWER FLOW MODELS WITH RETRAINING APPLICATIONS

William John Hayne

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

ANALYSIS OF SOME TWO-CHARACTERISTIC MARKOV-TYPE MANPOWER FLOW MODELS WITH RETRAINING APPLICATIONS

by

William J. Hayne

June 1974

Thesis Advisor: K. T. Marshall

Approved for public release; distribution unlimited.

T16 C41

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
ANALYSIS OF SOME TWO-CHARACTERISTIC
MARKOV-TYPE MANPOWER FLOW MODELS
WITH RETRAINING APPLICATIONS

5. TYPE OF REPORT & PERIOD COVERED

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
William J. Hayne

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93940

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93940

12. REPORT DATE
June 1974

13. NUMBER OF PAGES
154

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)
Naval Postgraduate School
Monterey, California 93940

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Manpower planning models
Manpower flow models
Markov models
Retraining models
Hierarchical organizations

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A discrete state, discrete time Markov-type manpower flow model having a two-dimensional state space (i.e., a two-characteristic model) is analyzed. The probabilistic properties of the model and the equations of stocks and flows are developed. A new method of representing the stocks as a sum of steady-state and transient components is presented. Two specific applications of the model to multi-grade hierarchical systems in which the

dimensions of the state spece are (grade, length of service) and (grade, time in grade) are analyzed in detail. The problem of combining states across the second dimension of the state space is studied and methods are derived for the (usual) case where the states are not lumpable. Finally, some applications of the model with state space (grade, skill group) to retraining problems between various skill categories are presented.

Analysis of Some Two-Characteristic Markov-Type Manpower Flow Models with Retraining Applications

by

William J. Hayne
Lieutenant Commander, United States Navy
B.S., Illinois Institute of Technology, 1961
M.S., Naval Postgraduate School, 1972

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

from the
Naval Postgraduate School
June 1974

Thesis
[illegible]
c.

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

ABSTRACT

A discrete state, discrete time Markov-type manpower flow model having a two-dimensional state space (i.e., a two characteristic model) is analyzed. The probabilistic properties of the model and the equations of stocks and flows are developed. A new method of representing the stocks as a sum of steady-state and transient components is presented. Two specific applications of the model to multi-grade hierarchical systems in which the dimensions of the state space are (grade, length of service) and (grade, time in grade) are analyzed in detail. The problem of combining states across the second dimension of the state space is studied and methods are derived for the (usual) case where the states are not lumpable. Finally, some applications of the model with state space (grade, skill group) to retraining problems between various skill categories are presented.

TABLE OF CONTENTS

I. INTRODUCTION -- 8

A. GENERAL DESCRIPTION --------------------------- 9

B. CONTENTS AND SUMMARY ------------------------- 12

II. PROBABILISTIC PROPERTIES OF THE TWO-CHARACTERISTIC MODEL ------------------------------ 18

A. DEFINITIONS AND ASSUMPTIONS ----------------- 19

B. FIRST-ORDER PROPERTIES ---------------------- 23

C. TWO SPECIAL CASES --------------------------- 28

D. VARIANCES ----------------------------------- 29

E. MATRICES OF t-STEP TRANSITION PROBABILITIES -- 30

F. CONDITIONING ON PROMOTION ------------------- 32

III. EQUATIONS OF STOCKS AND FLOWS ------------------- 37

A. DEFINITIONS AND BACKGROUND ------------------ 37

B. BASIC STOCK EQUATION ------------------------ 40

C. TRANSIENT PROPERTIES OF THE STOCKS ---------- 41

1. Fixed Exernal Flows --------------------- 46

2. Linear Growth of External Flows --------- 47

3. Geometric Growth of External Flows ------ 50

IV. THE (GRADE, LOS) MODEL ---------------------------- 53

A. BACKGROUND ---------------------------------- 53

B. DEFINITIONS AND DESCRIPTION OF THE MODEL ---- 55

C. COMPUTATIONAL CONSIDERATIONS ---------------- 61

D. EXAMPLE: MINIMIZATION OF COSTS OF REENLISTMENT ------------------------------- 62

E. INTERCHANGING GRADE AND LOS ----------------- 70

V. THE (GRADE, TIG) MODEL ------------------------- 75
  A. BACKGROUND ------------------------------- 75
  B. DESCRIPTION ------------------------------ 76
  C. INTERCHANGING GRADE AND TIG -------------- 80
VI. COMBINING STATES ------------------------- 83
  A. LUMPABILITY ------------------------------ 84
  B. LUMPING AN UNLUMPABLE PROCESS ------------ 86
    1. Constant Stocks ----------------------- 90
    2. Convergent Stock Distribution Vector --- 90
    3. Constant External Flows --------------- 91
    4. Linear Growth of External Flows ------- 95
  C. AN EXAMPLE ------------------------------- 96
VII. AN APPLICATION TO RETRAINING PROBLEMS --------- 101
  A. INTRODUCTION ----------------------------- 101
  B. DEFINITIONS AND THE BALANCE EQUATION ------ 104
  C. PARAMETRIC CALCULATIONS FOR PROMOTION RATES -- 106
  D. PARAMETRIC CALCULATIONS FOR STOCKS -------- 107
  E. RETRAINING FLOWS ------------------------- 110
  F. THE RETRAINING TRANSPORTATION PROBLEM ----- 111
  G. RELATION OF RETRAINING COSTS TO STOCKS ---- 115
    1. Special Case 1 -- Costs Constant ------ 121
    2. Special Case 2 -- Costs Increase with Stocks ------------------------------- 121
    3. Special Case 3 -- Costs Decrease with Stocks ------------------------------- 123
  H. TWO-GRADE OPTIMIZATION OF RETRAINING COSTS - 124
  I. THE RELATION OF TRAINING COSTS TO TOTAL FLOWS ------------------------------------ 127
APPENDIX A. Summary of Notation ------------------ 143
APPENDIX B. Diagonal Matrices -------------------- 145
LIST OF REFERENCES ------------------------------- 151
INITIAL DISTRIBUTION LIST ------------------------ 153

ACKNOWLEDGEMENTS

To Professor Kneale T. Marshall: Your experience, advice, and encouragement were invaluable in my research. Your caustic wit and unfailing patience made hard work an enjoyable project. Put briefly in Navy parlance, "Bravo Zulu."

To Professors Gordon Bradley, Toke Jayachandran, Paul Milch and Glen Myers: I sincerely appreciate your assistance. Your comments were most helpful in the presentation of the results.

To Professors James Esary, James Hartman, Gilbert Howard and P. A. W. Lewis: Sincere thanks for making my studies interesting and challenging.

To M. Wright who typed this thesis: My thanks for your patience and attention to detail.

To my wife, Jackie: You made it possible; you made it worth doing.

I. INTRODUCTION

This paper deals with manpower flow models, specifically discrete state, discrete time Markov-type models in which the states are described by two characteristics. This type of manpower flow model will be referred to as a "two-characteristic model" to distinguish it from the more common model in which the states are described by one characteristic.

Markov-type flow models have been used for some years; see for example Young and Almond (1961), Blumen, Kogan and McCarthy (1955), Marshall and Oliver (1970) and Rowland and Sovereign (1969). Theoretical properties of the model are developed in Bartholomew (1967). Charnes, Cooper and Niehaus (1972) use the Markov-type flow model as part of a larger manpower planning model.

The latter application is typical in that the Markov-type flow model is not used in isolation to project stocks and flows of people, but rather it is used to provide information to a more extensive model that deals with budgets, capital investments and aggregate work planning. The simple mathematical structure and the computational tractability of Markov-type models make it practical to embed them in larger models.

Early Markov-type models had one-characteristic states, such as pay grade or status in a hierarchical system. Geometric lifetimes in these states were thus implied by the model structure and often could not be justified. However,

the one-characteristic model usually led to a small state space which implied little computational difficulty. The simple mathematical structure of these models was also appealing, and analytical results were obtainable.

In an attempt to overcome the shortcomings of these simple models subsequent models in manpower planning used more than one characteristic to describe the states of the system. See for example the U.S. Navy ADSTAP models (U.S. Navy, 1973) and the U.S. Air Force TOPLINE model (RAND, 1973). In the documentation we have seen these have not been formally recognized as Markov-type models. Each model has been treated as an individual case. By viewing these models as special cases of a more general two-characteristic Markov-type model, we present in this thesis a unified treatment of such a model. We are able to develop the theory of its structure and we use this to analyze special cases frequently applied in military manpower planning. Some new applications of the model to the problem of retraining between skill groups in an organization are presented.

A. GENERAL DESCRIPTION

A graded manpower system is one in which each person is assigned a "grade," e.g., a pay grade, a rank or a status label. Such a system may be analyzed using a Markov-type model in which the states are defined to be the grades. Thus a state is identified by a single characteristic and we say the model has a one-dimensional state space. The analysis of a graded manpower system by such models is often suspect

because the grades do not provide a state space within which the system is actually Markovian. For example those in a given grade who progress rapidly through the lower grades may have significantly different promotion prospects from those in the same grade whose progress was less rapid. Thus the future progress of an individual in a state depends on how he reached the given state.

As is frequently the case in Markov models, expanding the state space may lead to a model of the system in which the Markov assumption is more realistic. In this paper the state space of a Markov-type model of a graded manpower system is expanded to one in which each state is identified by a couple (i,j) where i represents a grade and j represents a second characteristic. Examples of a second characteristic are:

1) Length of service in the system (LOS)
2) Time in grade (TIG)
3) Skill category
4) Overall performance mark
5) Physical location.

In cases 1) and 2) the second characteristic is chosen primarily to make the Markov assumption more tenable. In the other cases the appropriate second characteristic is chosen to fit the problem at hand.

Expanding the state space of a Markov-type model presents some difficulties, among which are:

1) There are more parameters (transition probabilities) in the model to be estimated or controlled.

This paper does not deal with the problems of parameter estimation. In practice the estimation of transition probabilities in manpower systems is hindered more by lack of data than by questions regarding appropriate statistical techniques. More pertinent is the number of parameters to be controlled, because the models analyzed here are intended for planning rather than forecasting. That is, we are interested in comparing the results of alternative policies rather than projecting the consequences of some "state of nature." Increasing the number of parameters to be controlled in a planning model can lead to practical difficulties. Such difficulties are not investigated in a comprehensive fashion in this paper; however, in Section D of Chapter IV and Sections C and D of Chapter VII we demonstrate methods which effectively reduce the number of parameters the planner must control.

2) Computations may be less tractable, if not completely impractical.

The manpower flow model is typically a component of a larger planning model, so it is to be expected that a single run of the planning model will require numerous calls for information from the flow model. Consider a system in which two grades i and k are important plateaus to be reached which affect an individual's retirement benefits. A retirement policy planning model might require information on the fraction of people who will attain grade k given that grade i is attained as well as the distribution of time required to reach grade k. Practical methods for deriving this type of information from a

two-characteristic model are developed in Chapters II and III. Although a two-characteristic model in a military manpower application would have at least one hundred states, it is shown that much information of interest can be derived by calculations restricted to a smaller number of states.

3) Analytical results which yield insight into the properties of the system may be more difficult to obtain.

Certainly analytical results are more difficult to obtain in a two-characteristic model than in a one-characteristic model. Nevertheless, we obtain a number of results which include the structure of the fundamental matrix (Chapter II), the equilibrium distribution of people in the system under various hiring policies (Chapter III), the effect of average time spent in a grade on the equilibrium distribution of people in that grade (Chapter VI) and the relation between promotion rates and the number in a grade (Chapter VII).

B. CONTENTS AND SUMMARY

Chapter II begins with the definition of basic notation and the statement of assumptions. Probabilistic properties of the transient matrix of the two-characteristic model, such as the expected number of visits to a state and the expected time spent in a grade are developed. A not uncommon property of manpower flow models, the 0-1 visiting property" (i.e., each state can be visited no more than once), is defined and its implications discussed. The "no return property" (i.e., a transition into a state is impossible once a transition has

been made out of it) is defined. If the states in a grade have the no return property then it may be assumed without loss of generality that the fundamental matrix for that grade is upper triangular and thus is relatively easy to compute. Matrices of the probabilities of going from one state to another in t steps (t-step transition matrices) are defined and a recursive formula for their calculation is shown. The t-step transition matrices are subsequently used in Chapter III to represent the stocks as the sum of steady-state and transient components. In the last section of Chapter II is it shown how the basic probabilistic properties of the transient matrix of the two-characteristic model are developed when one conditions on the attainment of a higher grade. The results of Chapter II follow closely the results of Kemeny and Snell (1960), but extend them to the two-characteristic model. Thus Chapter II contains a comprehensive treatment of how probabilistic information may be obtained from a two-characteristic model, and such a treatment has not been previously published.

Chapter III begins with a definition of the timing convention. In real systems people enter, leave and change grades or skills continuously. In order to fit a Markov-type model in discrete time the notion of a "period" must be carefully defined and a consistent counting convention used. Since planners almost invariably use discrete planning periods, these conventions lead to useable results. We assume time is divided into discrete periods (for example, years). Each period is denoted by the integer value of time

at the end of the period. People are counted only at the last instant of a period. Various stock and flow vectors are defined; the term stock refers to the number of people in a state or set of states at the end of a period, the term flow refers to the number of people who make transitions from one state in one period to another state in the next period. In Section B we develop equations for computing the expected stocks at time t from the values of stocks and flows prior to time t. Section C contains the development of a new method for expressing the stock vector as a sum of a "steady-state" component and a "transient" component. The steady-state component of the stock vector is derived for the cases where external flows are

1) Constant
2) Growing linearly,
3) Growing geometrically.

Chapter IV deals with the two-characteristic model where length of service (LOS) is the second characteristic. This model is called the (grade, LOS) model. The chapter begins with background material showing how practical considerations might lead the manpower planner to use the (grade, LOS) model. In Section B we present definitions and display some of the matrices associated with the (grade, LOS) model in order to show their rather simple structure. The concept of a cohort (a group of people who enter the system at the same time and in the same state) is discussed. It is noted that when entry to the system is restricted to a single state, the (grade, LOS) model preserves cohorts in the sense that members of

different cohorts never simultaneously occupy the same state. Section C discusses computationally efficient methods for performing the matrix multiplications needed to derive information from the (grade, LOS) model. The discussion in this section depends on Appendix B which treats the multiplication of "diagonal matrices." Section D shows how the structure of the (grade, LOS) model might be exploited in solving a problem regarding the minimization of the costs of reenlistment (in a military organization). In the last section of Chapter III, Section E, we discuss how the modeler may treat grade as the second characteristic and LOS as the first characteristic. Such an interchanged model is called the (LOS, grade) model. It is shown that the (LOS, grade) model is more flexible than the (grade, LOS) model.

Chapter V deals with the two-characteristic model where time in grade (TIG) is the second characteristic. This model is called the (grade, TIG) model. The chapter begins with a brief discussion of conditions under which the model might be appropriate. In Section B the definitions are stated and the structure of various matrices associated with the (grade, TIG) model is displayed and discussed. It is noted that the (grade, TIG) model is equivalent to a discrete semi-Markov process. In Section C we discuss how the modeler may treat grade as the second characteristic and TIG as the first characteristic. The resulting model, the (TIG, grade) model, is different in structure from the two-characteristic models previously studied. However it is briefly shown how one might derive information from such a model.

Chapter VI treats the problem of combining states in a two-characteristic model. The problem is of practical importance because often there is more than one reasonable choice of the second characteristic to consider, and these would lead to higher dimensional state spaces. For example, length of service, time in grade and skill group are three reasonable choices of the second characteristic in military manpower planning models. Only one of these can be chosen, so it is useful to have some guidance regarding how one might combine states across the characteristics that are not chosen for inclusion in the two-characteristic model. In Section A we investigate the condition under which the states in each grade of a two-characteristic model are "lumpable" as defined by Kemeny and Snell (1960). In Section B we discuss how one might combine states when the conditions for lumpability are not satisfied.

In Chapter VII the problem of retraining people between various skill groups is considered. Accordingly the second characteristic of our model is taken to be the individual's skill group. The retraining problem is formulated under a condition of equilibrium stocks and flows; attrition rates must be specified. The decision variables for the planner are stocks, promotion rates and total retraining costs. Of course these variables are interdependent. We treat stocks as the independent variables and show the relation of promotion rates to stocks when external flows are specified. Two methods for varying the stocks to achieve desired promotion rates are developed. Equations to compute the availabilities and

requirements for retrained people are shown. The total cost of retraining is modeled using the classical transportation problem. The foregoing availabilities and requirements are treated as the supplies and demands in the problem; the unit costs of retraining people are treated as the transportation costs. The relation of the total cost of retraining to the stocks (and promotion rates) is developed. We assume that the planner wants to minimize the total cost of retraining, however we do not assume that he can explicitly describe the constraints on stocks and promotion rates that must be observed in minimizing costs. We develop a technique in which the planner can vary the stocks in two successively indexed grades so as to reduce total retraining costs, while holding the stocks and promotion rates in all other grades unchanged. This technique is quite practical because the planner needs only to be able to recognize combinations of stocks and promotion rates that are acceptable rather than having to formally specify a set of constraints. The practicality of this technique is further enhanced by our demonstration that when hiring is restricted to the lowest pay grade the total retraining cost is a convex function of the numbers of people hired and promoted each period. Thus if the collection of acceptable combinations of stocks and promotion rates is a convex set any locally optimal solution to minimizing total retraining costs is globally optimal.

II. PROBABILISTIC PROPERTIES OF THE TWO-CHARACTERISTIC MODEL

In this chapter we develop various probabilistic properties of the two-characteristic model. One of our purposes here is to present a unified, computationally tractable approach to deriving probabilistic information from a two-characteristic model.

Let Q be the transition matrix for the transient states of a two-characteristic model. (The reader should see Chapter 3 of Kemeny and Snell (1960) for a general treatment of transient states in finite Markov chains.) For a typical military manpower planning application Q has dimensions at least 100 x 100. If the model were applied to the U.S. Navy enlisted force with nine pay grades as the first characteristic and 88 skill groups as the second characteristic, then Q would have dimensions 792 x 792. Military applications with even larger state spaces are not uncommon.

Following Kemeny and Snell (1960) we define the "fundamental matrix," denoted N, to be $(I-Q)^{-1}$. Many of the probabilistic properties of interest can be obtained from a knowledge of N. With large state spaces the computation of N is non-trivial. By imposing reasonable restrictions on the transitions between grades (no demotions and no multi-grade promotions) we are able to define a fundamental matrix N_i for each grade i, determine many probabilistic properties of the overall system from these comparatively small matrices, and show how these matrices are combined to form the fundamental

matrix N. In military manpower planning applications, the number of states in any grade is typically one fifth to one tenth of the total number of states in the system, so computing the fundamental matrix for a grade is significantly easier than computing the fundamental matrix for the entire system. Usually only limited portions of the fundamental matrix for the whole system are needed, and these may be readily calculated using the techniques developed here based on the fundamental matrices of the individual grades.

Certain matrix and vector notation used in this and subsequent chapters is summarized in Appendix A.

A. DEFINITIONS AND ASSUMPTIONS

Each state in the system is identified by a couple (i,j); a person is in state (i,j) if he is in grade i and has second characteristic j. There are n consecutively numbered grades: $i=1,\cdots,n$. Each grade i has a set of values of the second characteristic, denoted J(i), so the state (i,j) is defined for $j\varepsilon J(i)$ but not otherwise. For example if the second characteristic is length of service (LOS), and a person in grade 1 can be in this grade only when his LOS is between 1 and 3, then $J(1) = \{1,2,3\}$.

We assume that the second characteristic takes on consecutive integer values in each grade, so we define,

$l(i)$ = smallest value of the second characteristic associated with grade i,

$u(i)$ = largest value of the second characteristic associated with grade i.

Thus,

$$J(i) = \{j: \ j = l(i),\cdots,u(i)\}.$$

Let w_i be the number of elements in $J(i)$. Then,

$$w_i = u(i) - l(i) + 1.$$

Denote the set of states associated with grade i by T_i. Then,

$$T_i = \{(i,j): j \varepsilon J(i)\}, \ i=1,\cdots,n.$$

Note that T_i contains w_i elements.

Let $T = \bigcup_{i=1}^{n} T_i$, the complete set of all transient states, and let T_0 denote the single absorbing state "out of the system."

Let,

1) Q_i be a w_i x w_i matrix of one-step transition probabilities $q_i(j,m)$, $j,m \varepsilon J(i)$, where $q_i(j,m)$ is the probability of a transition from state $(i,j) \varepsilon T_i$ to state $(i,m) \varepsilon T_i$.

2) P_i be a w_i x w_{i+1} matrix of one-step transition probabilities $p_i(j,m)$, $j \varepsilon J(i)$, $m \varepsilon J(i+1)$, where $p_i(j,m)$ is the probability of a transition from state $(i,j) \varepsilon T_i$ to state $(i+1,m) \varepsilon T_{i+1}$.

3) A_i be a w_i x 1 matrix of one-step transition probabilities $a_i(j)$, $j \varepsilon J(i)$, where $a_i(j)$ is the probability of a transition from state (i,j) to "out of the system," i.e., to T_0.

The basic assumption in this thesis is

A0: Movement between states of the system follows the stochastic laws of a stationary finite state Markov chain.

The following restrictions on movements between grades are assumed:

A1: From any state in T_i it is possible to make a one-step transition only to states in T_i, T_{i+1} or T_0.

The practical significance of this assumption is that no person is promoted more than one grade in a single period, and no one is ever demoted. We have chosen for the sake of definiteness to call the first characteristic "grade." Any characteristic that satisfies A1 can be used as a first characteristic. Length of service is an example of a characteristic that satisfies A1, and we take advantage of this in Section E of Chapter IV.

We also assume:

A2: Each matrix Q_i, $i=1,\cdots,n$, is a transient matrix. The

The practical significance of this assumption is that no one can stay in a grade forever. With A1 this means that everyone entering the system must (with probability 1) eventually leave the system.

Under A0 through A2 the one step stochastic transition matrix for the entire system, denoted P, is

$$P = \left[\begin{array}{ccccc|c} Q_1 & P_1 & & & & A_1 \\ & Q_2 & P_2 & & & A_2 \\ & & Q_3 & \ddots & & A_3 \\ & & & \ddots & P_{n-1} & A_{n-1} \\ & & & & Q_n & A_n \\ \hline \multicolumn{5}{c|}{\bar{0}} & 1 \end{array}\right].$$

(Recall from Appendix A that $\bar{0}$ is a vector of zeroes.)

The transient matrix Q for the transient states T is then,

$$Q = \begin{bmatrix} Q_1 & P_1 & & & \\ & Q_2 & P_2 & & \\ & & Q_3 & \ddots & \\ & & & \ddots & P_{n-1} \\ & & & & Q_n \end{bmatrix}. \tag{1}$$

The plan of this chapter is to develop the probabilistic properties of:

1) any set of states T_i,

2) any union of consecutively indexed sets T_i,

i.e., $\bigcup_{i=k}^{m} T_i$,

3) the union of all transient states, T.

One of the purposes of this development is to show that the stochastic properties of Q, typically a large matrix, are readily calculated in terms of the smaller matrices Q_i and P_i.

The results of this chapter follow closely those in Chapter 3 in Kemeny and Snell (1960). The notation (K&S,3._._) indicates that a result can be derived from theorem 3._._ in Kemeny and Snell.

B. FIRST-ORDER PROPERTIES

The term "first-order properties" is used here as a general term for various probabilities and first moments, e.g., the probability of visiting a state, the expected number of visits to a state and the expected time spent in a grade.

Under A2 the overall transient matrix Q has a fundamental matrix $N = (I-Q)^{-1}$, and each element of N is the expected number of visits to the column state starting from the row state (K&S,3.2.4). Under A1, Q has the structure shown in (1). By direct methods it can be shown that

$$N = \begin{bmatrix} N_1 & N_1P_1N_2 & N_1P_1N_2P_2N_3 & \cdots & \prod_{i=1}^{n-1}(N_iP_i)N_n \\ & N_2 & N_2P_2N_3 & \cdots & \prod_{i=2}^{n-1}(N_iP_i)N_n \\ & & N_3 & \cdots & \prod_{i=3}^{n-1}(N_iP_i)N_n \\ & & & \ddots & \vdots \\ & & & & N_n \end{bmatrix} \qquad (2)$$

where $N_i=(I-Q_i)^{-1}$, $i=1,\cdots,n$, the fundamental matrix for grade i. Note that the large matrix N is completely determined by the matrices N_i and P_i. Thus the only matrix inversions required to find $(I-Q)^{-1}$ are the inversions of $(I-Q_i)$,

$i=1,\cdots,n$. This is of considerable computational significance because as previously mentioned, Q is usually a large matrix.

Each matrix N_i has a probabilistic interpretation. We pursue this interpretation and show that these matrices can be used to determine other probabilistic properties of interest.

In this and subsequent sections we make numerous definitions, the k^{th} definition is denoted by Dk.

Let us consider first the properties associated with a single set of states T_i and define:

D1. $v_i(j,m)$ = expected number of visits to state (i,m) given that grade i is entered in state (i,j)

D2. V_i = a $w_i \times w_i$ matrix having $v_i(j,m)$ as the element in row $j-l(i)+1$ and column $m-l(i)+1$.

From (2) the element of N_i in row $j-l(i)+1$ and column $m-l(i)+1$ equals the expected number of visits to state (i,m) given that grade i is entered in state (i,j). (K&S,3.2.4). So, from definitions D1 and D2, we have,

$$V_i = N_i. \tag{3}$$

Note that the rows and columns of N_i and V_i correspond to states in grade i in the same way as the rows and columns of Q_i.

Now define:

D3. $\tau_i(j)$ = expected time in grade i given that grade i is entered in state (i,j)

D4. $\tau_i = [\tau_i(l(i)),\cdots, \tau_i(u(i))]$, a $w_i \times 1$ vector.

The expected time spent in grade i equals the sum of the

expected number of visits to the various states in grade i. From (3) and D3,

$$\tau_i(j) = \text{component } (j-l(i)+1) \text{ of } N_i\bar{1},$$

and from D4,

$$\bar{\tau}_i = N_i\bar{1}, \text{ a } w_i \text{ x } 1 \text{ vector} \qquad (4)$$

(Recall from Appendix A that $\bar{1}$ is a vector with all components equal to one.)

We next turn our attention to where the process goes when it leaves grade i. From assumption A1 (no demotions, no multi-grade promotions) the process upon leaving T_i must enter either T_{i+1} or T_0. Next define:

D5. $b_i(j,m)$ = probability of entering grade i+1 in state (i+1,m) given that grade i is entered in state (i,j)

D6. B_i = a w_i x w_{i+1} matrix having $b_i(j,m)$ as the element in row $j-l(i)+1$ and column $m-l(i+1)+1$

D7. $b_i(j)$ = probability of ever entering T_{i+1} given that grade i is entered in state (i,j)

D8. $b_i = [b_i(l(i)),\cdots,b_i(u(i))]$, a w_i x 1 vector

D9. $b_{i0}(j)$ = probability of never entering T_{i+1} given that grade i is entered in state (i,j)

D10. $b_{i0} = [b_{i0}(l(i)),\cdots,b_{i0}(u(i))]$, a w_i x 1 vector.

From these definitions it follows that

$$B_i = N_iP_i, \text{ a } w_i \text{ x } w_i \text{ matrix (K\&S,3.5.4)}, \qquad (5)$$

$$b_i = B_i\bar{1}, \text{ a } w_i \text{ x } 1 \text{ vector},$$

and,

$b_{i0} = \bar{1} - b_i$

$= N_i A_i$, a $w_i \times 1$ vector.

The matrix B_i is particularly useful in our analyses. For example, let f_i be a $1 \times w_i$ vector of the number of people entering T_i. Then $f_i B_i$ is a $1 \times w_{i+1}$ vector of the number of these people who will eventually enter T_{i+1}. (K&S,3.3.6). This vector is used repeatedly in Chapter III.

Next we consider the first-order properties related to grades i and k where $i \leq k$. Define:

D11. $b((i,j),(k,m))$ = probability of entering grade k in state (k,m) given that grade i is entered in state (i,j)

D12. B_{ik} = a $w_i \times w_k$ matrix having $b((i,j),(k,m))$ as the element in row $j-l(i)+1$ and column $m-l(k)+1$.

From definitions D5 and D11 and a simple conditioning argument we have

$$b((i,j),(i+2,m)) = \sum_{r=l(i)}^{u(i)} b_i(j,r)\, b_{i+1}(r,m).$$

So, from D12,

$$B_{i,i+2} = B_i B_{i+2} .$$

Notice from D11 that B_{ii} is an identity matrix and from D5 that $B_{i,i+1} = B_i$. More generally it can be shown that for $i \leq k$,

$$B_{ik} = \sum_{r=i}^{k-1} B_r, \text{ a } w_i \times w_k \text{ matrix.}$$

Define:

D13. $v((i,j),(k,m))$ = expected number of visits to state

(k,m) given that grade i is entered in state (i,j).

D14. V_{ik} = a w_i x w_k matrix having v((i,j),(k,m)) as the element in row j-l(i)+1

D15. $b_{ik}(j)$ = probability of ever entering grade k, given that grade i is entered in state (i,j)

D16. $b_{ik} = [b_{ik}(l(i)),\cdots,b_{ik}(u(i))]$, a w_i x 1 vector.

Considering each row of B_{ik} as the part of an initial probability vector that applies to T_k, we then have,

$$V_{ik} = B_{ik}N_k, \text{ a } w_i \text{ x } w_k \text{ matrix (K\&S,3.5.4)}, \qquad (6)$$

and,

$$b_{ik} = B_{ik}\bar{1}, \text{ a } w_i \text{ x 1 vector.}$$

Define:

D17. $\tau_{ik}(j)$ = expected time in grade k given that grade i is entered in state (i,j)

D18. $\tau_{ik} = [\tau_{ik}(l(i)),\cdots, \tau_{ik}(u(i))]$, a w_i x 1 vector.

The expected time in a grade is the sum of the expected number of visits to states in that grade, so

$$\tau_{ik} = V_{ik}\bar{1}, \text{ a } w_i \text{ x 1 vector.}$$

This completes our study of the first-order properties related to the various grades of the system. The foregoing definitions by no means exhaust the first-order properties of the two-characteristic model that might conceivably be of interest. It is felt, however, that these properties will often be of practical interest and that other first-order properties may be readily derived from these.

C. TWO SPECIAL CASES

The elements of the fundamental matrix for grade i, N_i, have a somewhat different interpretation when the states in grade i have what we call the "0-1 visiting property." We say that a state has the 0-1 visiting property if the state can be visited no more than one time. Important examples of two-characteristic models in which all transient states have the 0-1 visiting property are the models in which the second characteristic is either length of service (see Chapter IV) or time in grade (see Chapter V).

If each state in T_i has the 0-1 visiting property, then the expected number of visits to a state in T_i is equal to the probability of visiting the state. The element of N_i in row j-l(i)+1 and column m-l(i)+1 may then be interpreted as the probability of visiting state (i,m) given that grade i is entered in state (i,j).

Another property of interest is the "no return property." we say that a state has the no return property if it is impossible to ever make a transition into the state after a transition has been made out of the state. The 0-1 visiting property implies the no return property, but they are not equivalent. For example, in modeling manpower flows in the U.S. Civil Service one might use "pay step" as a second characteristic. Each state is then a couple (grade, pay step). A person can stay in the same pay step for more than one period, so if there are no demotions then each state would have the no return property but not the 0-1 visiting property.

If the states in T_i have the no return property then it is possible to order the states in T_i so that Q_i is upper triangular. When Q_i is upper triangular so is $I-Q_i$ and the computation of the inverse of $I-Q_i$, i.e., the fundamental matrix for grade i, N_i, is considerably easier than in the general case.

If the states in T_i have the 0-1 visiting property, then not only is N_i upper triangular but also the elements of N_i on the main diagonal are all ones.

D. VARIANCES

The format in this section follows closely that of Section B, but here we are concerned with various second moment properties of the two-characteristic model.

Define:

D19. $v_{2,i}(j,m)$ = variance of the number of visits to state (i,m) given that grade i is entered in state (i,j)

D20. $V_{2,i}$ = a $w_i \times w_i$ matrix having $v_{2,i}(j,m)$ as the element in row $j-l(i)+1$ and column $m-l(i)+1$.

Following (K&S,3.3.3),

$$V_{2,i} = N_i(2(N_i)_{dg} - I) - (N_i)_{sq}.$$

(See Appendix A definitions of "dg" and "sq.")

Define:

D21. $\tau_{2,i}(j)$ = variance of the time spent in grade i given that grade i is entered in state (i,j)

D22. $\tau_{2,i} = [\tau_{2,i}(l(i)),\cdots,\tau_{2,i}(u(i))]$, a $w_i \times 1$ vector.

Following (K&S,3.3.5),

$$\tau_{2,i} = (2N_i - I)\tau_i - (\tau_i)_{sq} .$$

Define:

D23. $v_2((i,j),(k,m))$ = variance of the number of visits to state (k,m) given that grade i is entered in state (i,j)

D24. $V_2(i,k)$ = a w_i x w_k matrix having $v_2((i,j),(k,m))$ as the element in row j-l(i)+1 and column m-l(k)+1.

Following (K&S,3.3.6),

$$V_2(i,k) = V_{ik}(2(N_k)_{dg} - I) - (V_{ik})_{sq} .$$

Define:

D26. $\tau_2((i,j),k)$ = variance of time spent in grade k given that grade i is entered in state (i,j)

D27. $\tau_2(i,k) = [\tau_2((i,l(i)),k),\cdots, \tau_2((i,u(i)),k)]$. a w_i x 1 vector.

Following (K&S,3.3.6),

$$\tau_2(i,k) = B_{ik}(2N_k - I)\ \tau_k - (\tau_{ik})_{sq} .$$

If each state in T_i has the 0-1 visiting property, then the diagonal elements of N_i are equal to one, and,

$$(N_i)_{dg} = I,$$

$$V_{2,i} = N_i - (N_i)_{sq} ,$$

$$V_2(i,k) = V_{ik} - (V_{ik})_{sq} .$$

E. MATRICES OF t-STEP TRANSITION PROBABILITIES

In this section we consider the probability of being in state (k,m) t steps after being in state (i,j). The

matrices of these probabilities are called the t-step transition matrices. They are used in Chapter III to represent the stock vectors as a sum of steady-state and transient components. In Chapter IV the t-step transition matrices are shown to provide a description of how "cohorts" flow through the system.

Define:

D28. $m(t:(i,j),(k,m))$ = probability of being in state (k,m) t steps after being in state (i,j), $t = 0,1,2,\cdots$

D29. $M_{ik}(t)$ = a $w_i \times w_k$ matrix having $m(t:(i,j),(k,m))$ as the element in row $j-l(i)+1$ and column $m-l(k)+1$.

The rows of $M_{ik}(t)$ are associated with states in T_i; the columns of $M_{ik}(t)$ are associated with states in T_k.

We have immediately that

$$M_{ii}(0) = I.$$

From assumption A1 (no demotions, no multi-grade promotions) we have,

$$M_{ik}(t) = \bar{\bar{0}} \text{ if } i > k,$$

$$M_{ik}(t) = \bar{\bar{0}} \text{ if } t < k-i .$$

(Recall from Appendix A that $\bar{\bar{0}}$ denotes a matrix of zeroes.)

If the process is to be in state (k,m) exactly t steps after being in state (i,j), then it must be in some state in grades k or k-1 exactly t-1 steps after being in state (i,j). Conditioning on this fact leads to the recursive equation,

$$M_{ik}(t) = M_{ik}(t-1)Q_i + M_{i,k-1}(t-1)P_{i-1}, \quad t = 1,2,\cdots . \tag{7}$$

For any i and k the sum over t of the probability matrices $M_{ik}(t)$ gives the matrix of the expected number of visits to states in grade k starting from states in grade i. So we have,

$$\begin{aligned}\sum_{t=0}^{\infty} M_{ik}(t) &= V_{ik}\,, \quad i \le k, \qquad (8)\\ &= \bar{\bar{0}}\,, \quad \text{otherwise}\end{aligned}$$

By assumption A2 the Q_i matrices are transient, so V_{ik} is a matrix of finite elements. This implies that,

$$\lim_{t\to\infty} M_{ik}(t) = \bar{\bar{0}}. \qquad (9)$$

From (7) it can be shown by an inductive argument that

$$M_{ik}(t) = \sum_{r=0}^{t-1} M_{i,k-1}(t-1-r)P_{k-1}Q_k^r\,. \qquad (10)$$

The t-step transition matrices provide a rather comprehensive picture of how people move through a two-characteristic system.

F. CONDITIONING ON PROMOTION

In manpower planning one is often interested in conditional probabilities, e.g., the probability of attaining grade k given that grade i is attained. The stochastic properties of the transient matrix Q under conditioning on promotion are briefly developed in this section.

Define:

D30. (i,j:t) = the event "in state (i,j) at time t"

D31. T_k^* = the event "a transition is made into T_k before leaving the system."

Conditioning on the event T_k^* is the same as conditioning on promotion to grade k.

Define:

D32. $q_i(j,m) = \Pr[(i,m;t+1|(i,j;t)]$

D33. $q_i^*(j,m) = \Pr[(i,m;t+1)|(i,j;t), T_{i+1}^*]$

Provided that $\Pr[T_{i+1}^*|(i,j;t)] \neq 0$, we have

$$\begin{aligned} q_i^*(j,m) &= \Pr[(i,m;t+1)|(i,j;t), T_{i+1}^*] \\ &= \Pr[(i,m;t+1)|(i,j;t)] \\ &\qquad \times \frac{\Pr[T_{i+1}^*|(i,j;t), (i,m;t+1)]}{\Pr[T_{i+1}^*|(i,j;t)]} \\ &= q_i(j,m) \times \frac{\Pr[T_{i+1}^*|(i,m;t+1)]}{\Pr[T_{i+1}^*|(i,j;t)]} \\ &= q_i(j,m) \times \frac{b_i(m)}{b_i(j)} \end{aligned} \tag{11}$$

Define:

D34. C_i = a $w_i \times w_i$ matrix having the elements of b_i (see D8) on its main diagonal and zeroes elsewhere.

We will assume that promotion to grade i+1 is possible from every state in T_i. Under this assumption C_i^{-1} exists. If promotion to grade i+1 is impossible from some state (i,j) then we must avoid conditioning on an impossible event. This is readily accomplished by temporarily treating state (i,j) as part of T_0 (out of the system) and redefining J(i), Q_i, P_i and A_i accordingly.

Define:

D35. Q_i^* = a w_i x w_i matrix having $q_i^*(j,m)$ as the element in row $j-l(i)+1$ and column $m-l(i)+1$.

Then from (11) and D34,

$$Q_i^* = C_i^{-1} Q_i C_i \, .$$

The matrix Q_i^* is the matrix of within grade one-step transition probabilities conditioned on the attainment of grade i+1.

Define:

D36. $p_i(j,m) = \Pr[(i+1,m;t+1) \,|\, (i,j;t)]$

D37. $p_i^*(j,m) = \Pr[(i+1,m;t+1) \,|\, (i,j;t),\ T_{i+1}^*]$

D38. P_i^* = a w_i x w_{i+1} matrix having $P_i^*(j,m)$ as the element in row $j-l(i)+1$ and column $m-l(i+1)+1$.

We then have,

$$\begin{aligned} p_i^*(j,m) &= \Pr[(i+1,m;t+1) \,|\, (i,j;t),\ T_{i+1}^*] \\ &= \Pr[(i+1,m;t+1) \,|\, (i,j;t)] \\ &\quad \times \frac{\Pr[T_{i+1}^* \,|\, (i,j;t),(i+1,m;t+1)]}{\Pr[T_{i+1}^* \,|\, (i,j;t)]} \\ &= p_i(j,m) \times \frac{1}{b_i(j)} \, . \end{aligned}$$

Thus from D34 and D38,

$$P_i^* = C_i^{-1} P_i \, .$$

The matrix P_i^* is the matrix of one-step promotion probabilities conditioned on the attainment of grade i+1.

Because $(Q_i^*)^r = C_i^{-1} Q_i^r C_1$,

The fundamental matrix for grade i when we condition on promotion to grade i+1 is

$$N_i^* = (I-Q_i^*)^{-1}$$

$$= \sum_{r=0}^{\infty} (Q_i^*)^r$$

$$= \sum_{r=0}^{\infty} C_i^{-1} Q_i^r C_i$$

$$= C_i^{-1} N_i C_i .$$

Define:

D39. $v_i^*(j,m)$ = expected number of visits to state (i,m) given that grade i is entered in state (i,j) and grade i+1 is attained.

D40. V_i^* = a $w_i \times w_i$ matrix having $v_i^*(j,m)$ as the element in row j-l(i)+1 and column m-l(i)+1

D41. $b_i^*(j,m)$ = probability of entering grade (i+1) in state (i+1,m) given that grade i is entered in state (i,j) and grade i+1 is attained.

D42. B_i^* = a $w_i \times w_{i+1}$ matrix having $b_i^*(j,m)$ as the element in row j-l(i)+1 and column m-l(i+1)+1.

Then one may show that,

$$V_i^* = N_i^* ,$$

and,

$$B_i^* = N_i^* P_i^*$$

$$= C_i^{-1} B_i .$$

Note that B_i^* is simply B_i with its rows normalized; Q_i^* is not simply a row normalized form of Q_i.

As with the matrices B_i, products of matrices B_i^* with successive indices are well defined; their meaning is that of a matrix B_{ik} as defined in D11 and D12 with conditioning on attainment of grade k.

The conditioned and unconditioned matrices may be used together. For example, the elements of $B_i^* B_{i+1}$ give the probabilities of entering grade i+2 in the column state conditioned on starting from the row state in T_i and eventually attaining grade i+1.

III. EQUATIONS OF STOCKS AND FLOWS

We begin by defining the terms "stocks" and "flows" and then discuss why stocks and flows are important in manpower planning models. Next the relations between stocks and flows in a two-characteristic model are developed. Finally we show how the stocks can be represented as the sum of a "steady-state" component and a "transient" component.

A. DEFINITIONS AND BACKGROUND

The following three definitions constitute a "timing convention"; it is used in this and following chapters.

(1) A period is the interval of time from immediately after an integer value of the time parameter t up to and including the next integer value of t. A period is identified by the value of the time parameter at the end of the period. Thus,

$$\text{period } t_1 = \{t: \; t_1 - 1 < t \leq t_1\}$$

where t_1 is an integer.

(2) The number of people in a state at the end of a period is referred to as the "stock" in that state. Thus, stocks are counted only at integer values of the time parameter, t.

(3) The number of people who change their status in the system from one state to another during any period is referred to as a "flow." Flows occur during a period, but we do not specify the exact time at which they occur.

Stocks and flows are of primary importance in most manpower planning models. The most obvious reason for this is

that costs are closely related to stocks and flows, e.g., total payroll depends on stocks, transportation costs or retraining costs depend on flows. Recruiting policy and promotion policy depend in the short term on present stocks and in the long term on how we model future stocks and flows. Determining the feasibility of a retirement plan and evaluating the effects of a change in billet structure are other instances in which the planner needs to be able to model stocks and flows in a manpower system.

It should be noted that the Markov-type model is not the only method that one might use to model stocks and flows. A different method is the "cohort model." Marshall (1973) presents a comparison of the Markov-type and the cohort models. Another method for modeling stocks and flows is the "chain model" presented in Grinold and Marshall (to be published). A listing of various manpower flow models in U.S. Navy (1973) indicates that the Markov-type model is by far the most commonly used method in military applications.

We now define the variables that are used to model the stocks and flows in the two-characteristic model. Recall that T_i is the set of states associated with grade i, w_i is the number of states in T_i, and for convenience of notation we assume the second characteristic takes on successive integer values in grade i.

In a Markov model the stocks and flows are in general random variables. In this thesis we deal only with the expected values of stocks and flows. Such a model is called a "fractional flow model" because the transition probabilities

of the Markov model are in effect treated as fractions which direct flows through the system in a deterministic manner.

Let,

$s_{ij}(t)$ = expected stocks in state (i,j) at time t,

$s_i(t) = (s_{i,l(i)}(t),\cdots,s_{i,u(i)}(t))$,

a 1 x w_i vector of expected stocks in T_i,

$s(t) = (s_1(t),\cdots,s_n(t))$,

a 1 x $\sum_{i=1}^{n} w_i$ vector of expected stocks in the system.

By assumption A1, flows into any state in T_i must come from a state in either T_i or T_{i-1}. We will also make provision in our model for "external flows." The source of such flows is unspecified. However, we may consider external flows as consisting of people hired into the system. The external flows may be deterministic or random, but we deal only with their expected values.

Let,

$d_{ij}(t)$ = expected flow from states in T_i to state (i,j) during period t, a scalar:

$d_i(t) = (d_{i,l(i)}(t),\cdots, d_{i,u(i)}(t))$, a 1 x w_i vector;

$f_{ij}(t)$ = expected external flow into state (i,j) during period t, a scalar;

$f_i(t) = (f_{i,l(i)}(t),\cdots, f_{i,u(i)}(t))$, a 1 x w_i vector;

$g_{ij}(t)$ = expected flow from states in T_{i-1} to state (i,j) during period t, a scalar;

$$g_i(t) = (g_{i,1(i)}(t), \cdots, g_{i,u(i)}(t)), \text{ a } 1 \times w_i \text{ vector.}$$

When $i = 1$, $g_{ij}(t)$ is defined to be zero.

The relation between the flow vectors and the stock vector in grade i is depicted in Figure 1 where "T_i;t" denotes the states in grade i at time t.

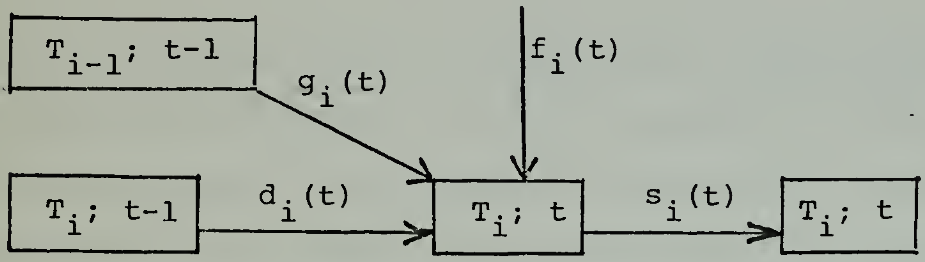

Figure 1. Stocks and Flows in Grade i in Period t.

B. BASIC STOCK EQUATION

By assumption A1 and the interpretation of external flows:

$$s_i(t) = d_i(t) + f_i(t) + g_i(t).$$

See Figure 1.

It will be convenient to define,

$$s_0(t) = \bar{0}, \text{ a vector of zeroes,}$$

$$P_0 = \bar{\bar{0}}, \text{ a matrix of zeroes.}$$

Using conditional expectation we then have

$$d_i(t) = s_i(t-1)Q_i, \quad i = 1, \cdots, n,$$

$$g_i(t) = s_{i-1}(t-1)P_{i-1}, \quad i = 1, \cdots, n.$$

The basic stock equation is then,

$$s_i(t) = s_i(t-1)Q_i + f_i(t) + s_{i-1}(t-1)P_{i-1}, \quad 1 \leq i \leq n. \quad (1)$$

The basic stock equation for grade i can be written in terms of the expected or actual stocks in grade i in previous

periods. By recursively applying the basic stock equation for $s_i(t)$, $s_i(t-1), \cdots, s_i(1)$ one obtains

$$s_i(t) = s_i(0)Q_i^{\,t} + \sum_{r=0}^{t-1} f_i(t-r)Q_i^{\,r} + \sum_{r=0}^{t-1} s_{i-1}(t-r-1)P_{i-1}Q_i^{\,r}, \qquad t = 0,1,2,\cdots, \quad i = 1,\cdots,n, \tag{2}$$

which we will refer to as the cumulative stock equation.

Equations (1) and (2) are used frequently in the remainder of this report. Some manpower models used in the U.S. military for short-range forecasting consist principally of an application of an equation similar to (1).

C. TRANSIENT PROPERTIES OF THE STOCKS

In this section we develop a method for expressing the stock vector as a sum of a "steady-state" component and a "transient" component. This method helps one to understand how the stock vectors will change in going from any present stock vector to future stock vectors. This method also helps one interpret the character of the limiting stock vector.

We do not want to restrict ourselves to cases in which the stock vector converges (as t increases) to a finite vector, so it is best to specify what is meant by a "steady-state" component of the stock vector. We say that the vector function $\tilde{s}_i(t)$ is the steady-state component of the stock vector $s_i(t)$ if,

$$\lim_{t\to\infty} (s_i(t) - \tilde{s}_i(t)) = \bar{0}\,.$$

For any sequence of stock vectors $\langle s_i(t)\rangle$ there is more than one choice of the steady-state component. In applications

one would prefer a steady state component having a relatively simple mathematical form. We show that in some cases a judicious choice of $\tilde{s}_i(0)$ makes this possible. The following theorem shows the properties of a class of steady-state components which can be quite useful.

Theorem. For any collection of 1 x w_i vectors $\tilde{s}_i(0)$, $i=1,\cdots,n$, let the vector functions $\tilde{s}_i(t)$ satisfy

$$\tilde{s}_i(t) = \tilde{s}_i(t-1)Q_i + f_i(t) + \tilde{s}_{i-1}(t-1)P_{i-1}, \quad t=1,2,\cdots, \quad i=1,\cdots,n.$$

i.e., the vector functions $\tilde{s}_i(t)$ satisfy the basic stock equation (1). Then.

(a) the actual stocks at time t are

$$s_i(t) = \tilde{s}_i(t) + \sum_{k=1}^{i} (s_k(0)-\tilde{s}_k(0))M_{ki}(t),$$

(b) $$\sum_{t=0}^{\infty} (s_i(t)-\tilde{s}_i(t)) = \sum_{k=1}^{i} (s_k(0)-\tilde{s}_k(0))B_{ki}N_i ,$$

a 1 x w_i vector having finite components,

(c) $\tilde{s}_i(t)$ is a steady-state component of the stock vector $s_i(t)$, i.e.,

$$\lim_{t\to\infty} (s_i(t)-\tilde{s}_i(t)) = \bar{0}.$$

Before proving the theorem it will be worthwhile explaining why one might be interested in such a theorem. Part (c) of the theorem says that $\tilde{s}_i(t)$ is a steady-state component of the stock vector $s_i(t)$, and part (a) shows how the stock vector $s_i(t)$ can be expressed as the sum of a steady-state component and a transient component. Part (b) of the theorem says that

the total over all periods of the difference between the stock vector and its steady-state component is a readily calculated finite vector.

Such information can be useful when long-range planning has been done using an "equilibrium model." As an example consider an organization which intends to change from its present size of 250,000 to a size of 200,000. The manpower planner may use an equilibrium model to develop policies that are in some sense optimal, and these policies will maintain the size of the organization at 200,000 people once it has been reduced to this size. So the equilibrium model tells the planner what to do once the size of the organization reaches the desired equilibrium level but it doesn't tell him how to change the size of the organization from its present level (250,000) to the desired equilibrium level (200,000). This problem of finding an optimal transition policy to go from present stock levels to a future equilibrium stock distribution is a difficult one. One method for making the transition is to immediately implement the hiring, promotion and attrition policies that have been derived from the equilibrium model. Because of the transient nature of the system (see assumption A2) these policies will eventually bring the stocks in the system to their equilibrium levels.

In the theorem the vector functions $\tilde{s}_i(t)$ play the role of what the stocks would be at time t if the system were in equilibrium. The stock vectors $s_i(t)$ indicate what the stocks will be at time t if we start with the present stocks $s_i(0)$ and implement the policies of the equilibrium model

(which are reflected in the external flows, $f_i(t)$, and the transition matrices Q_i, P_i and A_i). From part (a) of the theorem we may readily calculate the difference between actual stocks and equilibrium stocks in any grade and any period. If there is a penalty associated with having more people than the equilibrium stocks in the system, then part (b) of the theorem may be used to calculate the total penalty. Part (c) of the theorem assures the planner that the difference between the actual and equilibrium stocks does converge to a zero vector as the time parameter t increases.

The proof of the theorem follows.

Proof. By hypothesis the vector functions $\tilde{s}_i(t)$ satisfy the basic stock equation (1), so they must also satisfy the cumulative stock equation (2):

$$\tilde{s}_i(t) = \tilde{s}_i(0)Q_i^{\,t} + \sum_{r=0}^{t-1} f_i(t-r)Q_i^{\,r} + \sum_{r=0}^{t-1} \tilde{s}_{i-1}(t-r-1)P_{i-1}Q^{\,r} .$$

Of course the stock vectors $s_i(t)$ also satisfy the cumulative stock equation (2), so we have,

$$s_i(t)-\tilde{s}_i(t) = (s_i(0)-\tilde{s}_i(0))Q_i^{\,t} + \sum_{r=0}^{t-1} (s_{i-1}(t-r-1)-\tilde{s}_{i-1}(t-r-1))P_{i-1}Q_i^{\,r} .$$

When i=1 this implies,

$$\begin{aligned} s_1(t) &= \tilde{s}_1(t) + (s_1(0)-\tilde{s}_1(0))Q_1^{\,t} \\ &= \tilde{s}_1(t) + (s_1(0)-\tilde{s}_1(0))M_{11}(t), \end{aligned}$$

so we have shown that part (a) of the theorem is true when i=1. Suppose part (a) of the theorem is true for grade i-1, i.e.,

$$s_{i-1}(t) = \tilde{s}_{i-1}(t) + \sum_{k=1}^{i-1}(s_k(0)-\tilde{s}_k(0))M_{k,i-1}(t) .$$

Then

$$s_{i-1}(t-r-1)-\tilde{s}_{i-1}(t-r-1) = \sum_{k=1}^{i-1}(s_k(0)-\tilde{s}_k(0))M_{k,i-1}(t-r-1).$$

and,

$$\begin{aligned}
s_i(t)-\tilde{s}_i(t) &= (s_i(0)-\tilde{s}_i(0))Q_i^{\,t} \\
&\quad + \sum_{r=0}^{t-1}\sum_{k=1}^{i-1}(s_k(0)-\tilde{s}_k(0))M_{k,i-1}(t-r-1)P_{i-1}Q_i^{\,r} \\
&= (s_i(0)-\tilde{s}_i(0))Q_i^{\,t} \\
&\quad + \sum_{k=1}^{i-1}(s_k(0)-\tilde{s}_k(0))\sum_{r=0}^{t-1}M_{k,i-1}(t-r-1)P_{i-1}Q_i^{\,r}.
\end{aligned}$$

From Equation (10) in Section E of Chapter II,

$$\sum_{r=0}^{t-1}M_{k,i-1}(t-r-1)P_{i-1}\,Q_i^{\,r} = M_{ki}(t) ,$$

so we have shown by induction that,

$$s_i(t)-\tilde{s}_i(t) = (s_i(0)-\tilde{s}_i(0))Q_i^{\,t} + \sum_{k=1}^{i-1}(s_k(0)-\tilde{s}_k(0))M_{ki}(t) .$$

This proves part (a) of the theorem.

From part (a),

$$s_i(t)-\tilde{s}_i(t) = \sum_{k=1}^{i}(s_k(0)-\tilde{s}_k(0))M_{ki}(t) ,$$

$$\begin{aligned}
\sum_{t=0}^{\infty}(s_i(t)-\tilde{s}_i(t)) &= \sum_{t=0}^{\infty}\sum_{k=1}(s_k(0)-\tilde{s}_k(0))M_{ki}(t) \\
&= \sum_{k=1}^{i}(s_k(0)-\tilde{s}_k(0)\sum_{t=0}^{\infty}M_{ki}(t) \\
&= \sum_{k=1}^{i}(s_k(0)-\tilde{s}_k(0))B_{ki}N_i ,
\end{aligned}$$

a 1 x w_i vector having finite components.

The last step above follows from equations (6) and (8) of Chapter II. This proves part (b) of the theorem.

Part (c) follows from the fact that the sum in part (b) is finite. $\square$

The utility of this approach depends on our ability to find vectors $\tilde{s}_i(0)$ such that the vector functions $\tilde{s}_i(t)$ are simple and readily calculated. The following subsections present examples.

1. Fixed External Flows

The equilibrium models previously mentioned enjoy some popularity in military manpower planning in the United States. The rationale underlying the use of such models is that one should determine the organization structure and the policies to maintain this structure which are optimal. Among the policies derived from an equilibrium model is the hiring policy. This has the form,

$$f_i(t) = f_i, \quad t = 1,2,\cdots, \qquad i = 1,\cdots,n,$$

where the vector of the number of people to be hired into the states in grade i each period, f_i, is specified from the equilibrium model.

Define,

$$\tilde{s}_1(0) = f_1 N_1 .$$

Then using (1) it is easy to show that

$$\tilde{s}_1(t) = f_1 N_1 \text{ for all } t.$$

Thus, from the theorem

$$s_1(t) = \tilde{s}_1(t) + (s_1(0) - \tilde{s}_1(0)) M_{11}(t)$$

$$= f_1 N_1 + (s_1(0) - f_1 N_1) Q_1^{\,t} .$$

Now recursively define,

$$\tilde{s}_1 = \tilde{s}_1(t) = f_1 N_1 ,$$

$$\tilde{s}_i = (f_i + \tilde{s}_{i-1} P_{i-1}) N_i , \quad i = 2,\cdots,n . \tag{3}$$

It is straightforward to verify that these $\tilde{s}_i$ satisfy the basic stock equation (1), so we have from the theorem, when $f_i(t) = f_i$,

$$s_i(t) = \tilde{s}_i + \sum_{k=1}^{i} (s_k(0) - \tilde{s}_k) M_{ki}(t) .$$

The steady-state component can also be written,

$$\tilde{s}_i = \sum_{k=1}^{i} f_k B_{ki} N_i, \quad i = 1,\cdots,n . \tag{4}$$

Note that $\sum_{k=1}^{i} f_k B_{ki}$ is a nonnegative 1 x w_i vector, so the limiting vector of stocks in grade i must be a nonnegative combination of the rows of N_i. In general, then, not all nonnegative 1 x w_i vectors are possible limiting stock vectors under constant external flows.

2. Linear Growth of External Flows

In this section we consider the case in which the number of people hired into each state increases by the same amount each period. Such a hiring policy may not be natural over a long period of time, but it may provide a simple approximation to planned hiring policies.

Let the 1 x w_i vector f_i be the amount that the number hired into states in grade i increases each period. Then the external flow vector for grade i is,

$$f_i(t) = t f_i, \quad t = 1,2,\cdots, \quad i = 1,\cdots,n .$$

Let,

$$\tilde{s}_1(0) = f_1 N_1 Q_1 N_1 .$$

Let the vector function $\tilde{s}_1(t)$ satisfy the basic stock equation (1),

$$\tilde{s}_1(t) = \tilde{s}_1(t-1) Q_1 + f_1(t) .$$

Using the identity $N_1 Q_1 + I = N_1$ one can show that

$$\tilde{s}_1(t) = t f_1 N_1 - N_1 Q_1 N_1 .$$

Thus from the theorem,

$$s_1(t) = t f_1 N_1 - f_1 N_1 Q_1 N_1 + (s_1(0) + f_1 N_1 Q_1 N_1) Q_1^{t} .$$

We note that $s_1(t)$ is of the form

$$\tilde{s}_1(t) = t L_1 + C_1$$

where $L_1 = f_1 N_1$ is a 1 x w_i vector,

and $C_1 = -f_1 N_1 Q_1 N_1$ is a 1 x w_i vector.

Consider some grade $i \varepsilon \{2, \cdots, n\}$. Suppose that

$$\tilde{s}_{i-1}(t) = t L_{i-1} + C_{i-1} ,$$

where L_{i-1} and C_{i-1} are 1 x w_{i-1} vectors.

Using the identify

$$(t f_i N_i - f_i N_i Q_i N_i) Q_i + (t+1) f_i = ((t+1) f_i N_i - f_i N_i Q_i N_i) ,$$

one may show that if

$$\tilde{s}_i(t) = t f_i N_i - f_i N_i Q_i N_i + \tilde{s}_{i-1}(t-1) P_{i-1} N_i - L_{i-1} P_{i-1} N_i Q_i N_i ,$$

then $\tilde{s}_i(t)$ satisfies the basic stock equation (1). Note that $\tilde{s}_i(t)$ has the form,

$$\tilde{s}_i(t) = t L_i + C_i , \tag{5}$$

where,

$$L_i = f_iN_i + L_{i-1}P_{i-1}N_i$$

$$= (f_i + L_{i-1}P_{i-1})N_i \,, \qquad (6)$$

and

$$C = -(f_i + L_{i-1}P_{i-1})N_iQ_iN_i - (L_{i-1} - C_{i-1})P_{i-1}N_i$$

$$= -((L_{i-1} - C_{i-1})P_{i-1} + f_iN_iQ_i)N_i \,.$$

Thus we have shown that when the external flows grow linearly the steady-state component of the stocks also grows linearly.

By recursive substitution in (6) we have,

$$L_i = \sum_{k=1}^{i} f_kB_{ki}N_i \,.$$

Note that this vector gives the expected number of visits to states in grade i of $f_k=f_k(t+1)-f_k(t)$ entrants in grade k, $k=1,\cdots,i$. That is, the *growth* in the stocks in grade i each period, L_i, equals the expected number of visits to grade i of the *growth* in the external flows each period in the grades less than or equal to i.

Both L_i and C_i have the fundamental matrix N_i as a right factor, so the steady state component of the stock vector, $\tilde{s}_i(t)$, must be a nonnegative combination of the rows of N_i. This same result was observed in the case of constant external flows.

In summary we have shown that by choosing

$$\tilde{s}_i(t) = tL_i + C_i$$

where $L_i = f_1N_1$ when $i=1$,

$$= (f_i + L_{i-1}P_{i-1})N_i \,, \quad i=2,\cdots,n \,.$$

and $C_i = -f_1N_1Q_1N_1$ when $i=1$,

$$= -((L_{i-1} - C_{i-1})P_{i-1} + f_iN_iQ_i)N_i \ , \quad i=2,\cdots,n \ ,$$

then from the theorem the stock equation may be written

$$s_i(t) = \tilde{s}_i(t) + \sum_{k=1}^{i} (s_k(0)-\tilde{s}_k(0))M_{ki}(t) \ .$$

3. Geometric Growth of External Flows

In this subsection we will show that geometric growth of external flows leads (eventually) to geometric growth of the stocks. Geometric growth is a not uncommon phenomenon both in natural and man-made systems. Geometric growth is frequently a reasonable assumption for medium- to long-range planning in manpower systems.

We consider the case in which the external flows into the states in grade i are proportional to a known vector f_i and grow geometrically at a rate θ_i. Thus,

$$f_i(t) = \theta_i{}^t f_i, \quad t = 1,2,\cdots,$$
$$i = 1,\cdots,n$$
$$\theta_i > 0.$$

When $0 < \theta_i < 1$, the external flows contract rather than grow.

If θ_k is not an eigenvalue of Q_i for $k \le i \le n$ we may define,

$$N_i(\theta_k) = (I - \frac{1}{\theta_k} Q_i)^{-1} \ .$$

If the states in grade i have the 0-1 visiting property then all eigenvalues of Q_i are zero and thus $\theta_k > 0$ is never equal to an eigenvalue of Q_i in this case.

The following identity will be useful:

$$N_i(\theta_k) = \sum_{r=0}^{\infty} (\frac{1}{\theta_k} Q_i)^r .$$

From this it follows that

$$N_i(\theta_k)Q_i = \theta_k(-I+N_i(\theta_k)).$$

Define,

$$\tilde{s}_1(0) = f_1N_1(\theta_1) .$$

Then it can be shown that if

$$\tilde{s}_1(t) = \theta_1^t f_1N_1(\theta_1) .$$

then $\tilde{s}_1(t)$, $t=0,1,\cdots$, satisfies the basic stock equation, and from the theorem,

$$s_1(t) = \theta_1^t f_1N_1(\theta_1) + (s_1(0)-f_1N_1(\theta_1))M_{11}(t).$$

Note that the steady state component of the grade 1 stock vector grows geometrically at the same rate as the external flows into grade 1.

Define,

$$B_{ki}(\theta_k) = \prod_{m=k}^{i-1} (N_m(\theta_k)P_m), \quad 1 \leq k \leq i \leq n.$$

Then it can be shown that if

$$\tilde{s}_i(t) = \sum_{k=1}^{i} \theta_k^{t-(i-k)} f_kB_{ki}(\theta_k)N_i(\theta_k)$$

then $\tilde{s}_i(t)$, $t=0,1,\cdots$, satisfies the basic stock equation (1). Note that in the limit the stocks in grade i grow geometrically at the rate of the largest θ_k where $k \leq i$.

Define,

$$\theta_M = \max\{\theta_k; k=1,\cdots,i\}.$$

The steady-state component of the stock vector is not in general a nonnegative combination of the rows of N_i (as was the case with constant external flows and linear growth of external flows). Rather the steady-state stock distribution is a nonnegative combination of the rows of $N_i(\theta_M)$. The rows of $N_i(\theta_M)$ need not be nonnegative combinations of the rows of N_i, so the limiting stock distributions that are possible under geometric growth of external flows need not be the same as the limiting stock distributions under constant external flows and linear growth of external flows.

IV. THE (GRADE, LOS) MODEL

The (grade, LOS) model is a model of a graded manpower system in which the second characteristic is length of service (LOS). By a person's length of service we mean the number of periods that he has been in the system.

A. BACKGROUND

Various manpower planning models that account for both grade and LOS are presently used by the United States military services. The incorporation of length of service into military manpower flow models is important in order to realistically analyze the policies of enlisted contracts and "retirement at 20."

A person enlisting in the military service "signs on" for a number of years, usually two to five years. In the past, approximately eighty percent of those entering enlisted military service left the system upon expiration of their initial contract. This high attrition at initial contract expiration has had many effects, but two that are pertinent here are:

1) There is a close relation between attrition and length of service for men serving on their initial contract.

2) At least half the people in the system are serving under their initial contract, so it is generally quite important that this group be modeled accurately.

The inclination of the operations researcher may then be to model manpower flows by a Markov-type model in which the states are lengths of service. Such models are structurally simple and computationally efficient. The author has worked on interactive retraining models of this sort for the U.S. Marine Corps. See also Grinold, Marshall and Oliver, 1973. But in practice manpower flow models are typically only part, albeit a crucial part, of larger planning models. Budget planners and operational planners usually demand manpower projections aggregated by pay grade; the distribution of people by their length of service is usually of secondary interest.

Faced with the demand for a model that aggregates by pay grade and knowing that a valid model must treat the effects of length of service the operations researcher is led to conclude that a (grade, LOS) type of model is appropriate.

The foregoing discussion does not imply that the (grade, LOS) model is appropriate only for military organizations. Various members of the "English school" of manpower planning have investigated the appropriate distribution function for the length of time that a worker stays with a company [Silcock, 1954; Lane and Andrew, 1955; Bartholomew, 1959]. The constant failure rate distributions (exponential/ geometric) were found to be quite inadequate, indicating that attrition rate and length of service are indeed related.

B. DEFINITIONS AND DESCRIPTION OF THE MODEL

Let a person's LOS be the number of times the person has been counted in the system. Recall that under our timing convention a person who enters the system during period t is counted for the first time at the end of that period (at time t), and is assigned an LOS of 1 at that time. A person's LOS increases by one for each successive end-of-period that he is counted in the system.

It is assumed that once a person leaves the system he never returns. It is possible to modify the (grade, LOS) model to allow for departure and re-entry by assigning to each grade dummy states in which the LOS remains constant from one period to the next. In the model described here the "out of the system" state is treated as absorbing.

The states of the system are defined by couples (i,j) where:

(i,j) = the state corresponding to grade i and LOS j.

The notation and results for the general two-characteristic model apply directly to the LOS model. In particular

l(i) and u(i) are the lower and upper lengths of service for anyone in grade i.

By definition as long as a person remains in the system his LOS must increase by one each period. (This is an example of a model in which all transient states have the 0-1 visiting property.) Consequently, we need define only the following transition probabilities:

q_{ij} = probability a person in state (i,j) at the end

of one period will be in state (i,j+1) at the end of the next period,

p_{ij} = probability a person in state (i,j) at the end of one period will be in state (i+1,j+1) at the end of the next period,

a_{ij} = probability a person in state (i,j) at the end of one period will be out of the system at the end of the next period.

By assumption A1:

$$q_{ij} + p_{ij} + a_{ij} = 1.$$

The transition matrix A_i is a $w_i \times 1$ matrix:

$$A_i = [a_{i,l(i)},\cdots,a_{i,u(i)}].$$

The transition matrix Q_i is $w_i \times w_i$ and has non-zero elements only immediately above the main diagonal:

$$Q_i = \begin{bmatrix} 0 & q_{i,l(i)} & & & \\ & 0 & q_{i,l(i)+1} & & \\ & & 0 & \ddots & \\ & & & \ddots & q_{i,u(i)-1} \\ & & & & 0 \end{bmatrix}. \tag{1}$$

The transition matrix P_i is $w_i \times w_{i+1}$ and has non-zero elements only on a single diagonal band. If $l(i+1) \geq l(i)+1$ and $u(i+1) \geq u(i)+1$, then P_i has the form shown below, where:

1) the first max $\{0,l(i+1)-(l(i)+1)\}$ rows are zeroes

2) the last max $\{0,u(i+1)-(u(i)+1)\}$ columns are zeroes.

Thus,

$$P_i = \begin{bmatrix} 0 & & & & & & \\ 0 & & & & & & \\ \vdots & & & & & & \\ p_{i,l(i+1)-1} & & & & & & \\ 0 & p_{i,l(i+1)} & & & & & \\ & 0 & \ddots & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & & 0 & p_{i,u(i)} & 0 \cdots 0 \end{bmatrix} . \tag{2}$$

If $l(i+1) \leq l(i)$, the first $l(i)+1-l(i+1)$ columns of P_i are zeroes. If $u(i+1) \leq u(i)$, the last $u(i)+1-u(i+1)$ rows of P_i are zeroes. Under any circumstances P_i is a "diagonal matrix," (see Appendix B).

Define,

$$\begin{aligned} q_i(r,r+k) &= \sum_{j=r}^{r+k-1} q_{ij} \quad \text{if } k \geq 1, \\ &= 1 \quad \text{if } k=0 \\ &= 0 \quad \text{if } k<0. \end{aligned}$$

Then for $k=0,1,\cdots,w_i-1$ we have

$$Q_i^k = \begin{bmatrix} 0 \ldots 0 & q_i(l,l+k) & & & \\ & 0 & q_i(l+1,l+1+k & & \\ & & \ddots & \ddots & \\ & & & \ddots & q_i(u-k,u) \\ & & & & 0 \\ & & & & 0 \end{bmatrix},$$

where $l = l(i)$,

$u = u(i)$,

and the first k columns and the last k rows are zeroes.

When $k \geq w_i$, $Q_i^k = \bar{\bar{0}}$. Recalling that

$$N_1 = (I - Q_i)^{-1}$$

$$= \sum_{k=0}^{\infty} Q_i^k,$$

we have in the LOS model,

$$N_i = \sum_{k=0}^{w_i - 1} Q_i^k$$

$$= \left[\begin{array}{ccccc} 1 & q_i(l,l+1 & q_i(l,l+2) & \cdots & q_i(l,u) \\ & 1 & q_i(l+1,l+2) & \cdots & q_i(l+1,u) \\ & & & & q_i(u-1,u) \\ & & & & 1 \end{array}\right]. \qquad (3)$$

The notation hides the rather simple structure of the fundamental matrix N_i in the (grade, LOS) model. Consider the case where $l(i)=1$ and $u(i)=4$. We then have for grade i,

$$N_i = \left[\begin{array}{cccc} 1 & q_{i1} & q_{i1}q_{i2} & q_{i1}q_{i2}q_{i3} \\ & 1 & q_{i2} & q_{i2}q_{i3} \\ & & 1 & q_{i3} \\ & & & 1 \end{array}\right]. \qquad (4)$$

It is interesting to compare the structure of N_i in the (grade, LOS) model with the structure of a matrix, denoted B, of the submatrices B_{ik} from the general two-characteristic model:

$$B = \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ & B_{22} & B_{23} & B_{24} \\ & & B_{33} & B_{34} \\ & & & B_{44} \end{bmatrix}$$

From Section B of Chapter II,

$$B_{ik} = \prod_{r=i}^{k-1} B_r, \text{ a } w_i \times w_k \text{ matrix },$$

so the matrix B can be written,

$$B = \begin{bmatrix} I & B_1 & B_1B_2 & B_1B_2B_3 \\ & I & B_2 & B_2B_3 \\ & & I & B_3 \\ & & & I \end{bmatrix} . \qquad (5)$$

Recall that each element of B_i gives the probability of entering grade i+1 in a particular state given the state in which grade i was entered. Thus, the matrices B_i summarize information about transitions from one grade to the next just as the q_{ij}'s summarize information about transitions from one LOS to the next. In Chapter VI we consider combining states in a two-characteristic model; the foregoing discussion indicates that if all states in a grade are combined then the information originally summarized in the $w_i \times w_{i+1}$ matrix B_i must now be summarized by a single number, i.e. a 1 x 1 matrix B_i in the model that results from combining states.

In manpower modeling the term cohort usually refers to a group of people who enter the system at the same time and in

the same state. For example the group of freshmen entering a college in a particular year comprise a cohort. After the cohort has entered the system we use the term cohort to refer to the members who are still in the system.

Let us consider the movements of cohorts in the (grade, LOS) model when entry to the system is restricted to state (1,1). The cohort which enters during period t_1 must, at any subsequent time $t_1 + t$, be in the set of states (i,j) such that $j = t + 1$. Furthermore any cohort which entered the system during period $t_2 \neq t_1$ cannot at time $t_1 + t$ be in any of the states $(i, t+1)$. Thus when entry to the system is restricted to a particular state, e.g., state (1,1), the (grade, LOS) model preserves cohorts in the sense that members of different cohorts never simultaneously occupy the same state. The motion of a cohort through the system is clearly described by the t-step transition matrices.

Recall that the t-step transition matrix from states in grade k to states in grade i, $M_{ki}(t)$, has elements:

$M(t; (k,m), (i,j)$ = probability of being in state (i,j)
t steps after being in state (k,m).

The row index of $M_{ki}(t)$ is $r = m - l(k) + 1$; the column index of $M_{ki}(t)$ is $c = j - l(i) + 1$. Because LOS increases by one each period, if

$$M(t; (k,m), (i,j)) \neq 0,$$

then $j - m = t$.

But $j - m = t$ implies that

$$(c + l(i) - 1) - (r + l(k) - 1) = t, \text{ or}$$

$$c - r = t + (l(k) - l(i)).$$

Thus the non-zero elements of M_{ki} are in row r and column c such that

$$c - r = t + (l(k) - l(i)). \qquad (6)$$

so $M_{ki}(t)$ is a "diagonal matrix" with index $t+(l(k)-l(i))$. (See Appendix B.) Furthermore the non-zero diagonal of $M_{ki}(t+1)$ must be immediately above the non-zero diagonal of $M_{ki}(t)$. Recalling that,

$$\sum_{t=0}^{\infty} M_{ki}(t) = V_{ki}, \text{ (see equation (8) of Chapter II),}$$

we see that the non-zero diagonal of $M_{ki}(t)$ must be the set of elements of V_{ki} in row r and column c satisfying (6). Thus in the (grade, LOS) model all t-step transition matrices can be readily derived from the matrices V_{ki}.

C. COMPUTATIONAL CONSIDERATIONS

Because of the sparseness of the Q_i and P_i matrices it would be most inefficient to store and manipulate the entire matrices. In Appendix B the storage and multiplication of "diagonal matrices" is discussed. A "diagonal matrix" is defined as any matrix (not necessarily square) having the property that for some k the element in row i and column j equals zero if $i-j \neq k$. The matrices Q_i and P_i in the (grade, LOS) model are diagonal matrices.

It is shown in Appendix B that the product of diagonal matrices is a diagonal matrix. Only the elements of the non-zero diagonal and four numbers describing the diagonal matrix need be stored. To multiply an $r \times m$ diagonal matrix by an $m \times c$ diagonal matrix requires $\min\{r,m,c\}$ multiplications

as opposed to the product of r,m and c multiplications and additions in the general case. Consequently, in the (grade, LOS) model the fundamental matrix N_i can be computed efficiently by computing the non-zero diagonals of successive powers of Q_i and combining these with an identity matrix to form the fundamental matrix N_i.

D. EXAMPLE: MINIMIZATION OF COSTS AT REENLISTMENT

In this section we consider an example in order to illustrate how one might take advantage of the highly structured form of the (grade, LOS) model to solve a practical problem.

The U.S. Navy has used various reenlistment bonus plans to decrease the attrition rate at termination of the initial enlistment contract. The amount and effectiveness of the bonus depends on a number of factors; among them is pay grade at contract termination.

We consider the case in which initial contracts are for m periods, and a reenlistment bonus is used to control the attrition rate a_{km} for such grades k that a_{km} is defined. To simplify notation it is assumed that a_{km} is defined for $k=1,\cdots,n$. The cost of changing a_{km} is specified by a non-negative quasi-convex function. Let $c_k(\Delta a_{km})$ be the cost of a Δa_{km} decrease in a_{km} from its base value, and let $c_k(0) = 0$. Decisions are constrained by the requirements that in equilibrium the total stocks in each grade, $s_i\bar{1}$, must be no smaller than a specified lower bound s_i^-. All external flow is into grade 1 and f_1 is specified.

Because under A1,

$p_{km} + q_{km} + a_{km} = 1,$

if a_{km} is decreased then changes in either p_{km} or q_{km} must occur.

Let,

Δq_{km} = the increase in q_{km},

Δp_{km} = the increase in p_{km}.

Then,

$$\Delta q_{km} + \Delta p_{km} = \Delta a_{km} \,. \tag{7}$$

The various transition matrices depend on the values of the Δq_{km}'s and Δp_{km}'s, and we denote this by $Q_k(\Delta q_{km})$, $N_k(\Delta q_{km})$, $P_k(\Delta p_{km})$ and $B_k(\Delta q_{km})$. To simplify notation somewhat we define,

$Q_k = Q_k(0),$

$N_k = N_k(0),$

$P_k = P_k(0),$

$B_k = B_k(0,0).$

The parameters of the matrices Q_k and P_k are known, and N_k and B_k are calculated from them.

From the results of Chapter III we have

$$s_i = f_1 \prod_{r=1}^{i-1} B_r(\Delta q_{rm}, \Delta P_{rm}) N_i(\Delta q_{im}) \,. \tag{8}$$

In the case $\Delta q_{km} = \Delta p_{km} = \Delta a_{km} = 0$ for all k we denote the stocks by s_{i0} and,

$$s_{i0} = f_1 \prod_{r=1}^{i-1} B_r N_i \,. \tag{9}$$

To minimize the total costs of the reenlistment bonuses while satisfying the constraint that total stocks in grade i be no smaller than s_i^- we must solve the following program denoted P1.

$$\text{P1]} \quad \min \sum_{k=1}^{n} c_k(\Delta a_{km})$$

$$\text{ST } s_i\bar{1} \geq s_i^-, \quad i = 1,\cdots,n$$

$$s_i = f_1 \prod_{r=1}^{i-1} B_r(\Delta q_{rm}, \; \Delta p_{rm}) N_i(\Delta q_{im}), \quad i = 1,\cdots,n$$

$$\Delta q_{km} + \Delta p_{km} = \Delta a_{km}, \quad k = 1,\cdots,n$$

It will be shown that because of the structure of the (grade, LOS) model the constraints imposed by equation (1) are linear functions of the decision variables in program P1. That is, program P1 requires the minimization of a quasi-convex function subject to linear constraints.

The linearity of the constraints is shown in two steps. First, we show that they are linear when only one of the decision variables is non-zero. Next, we show they are linear for all feasible values of the decision variables.

When $\Delta q_{km}=\Delta p_{km}=\Delta a_{km}=0$ we denote the values of the parameters q_{km}, p_{km} and a_{km} with hats, i.e.,

$$q_{km} = \hat{q}_{km} + \Delta q_{km},$$

$$p_{km} = \hat{p}_{km} + \Delta p_{km},$$

$$a_{km} = \hat{a}_{km} - \Delta a_{km},$$

The value of s_i depends on q_{km} only through $N_k(\Delta q_{km})$. Let,

$$j = m - l(k) + 1 .$$

Then the only elements of $N_k(\Delta q_{km})$ that are functions of q_{km} are those in row r and column c such that $r \leq j$ and $c > j$. Each of these elements has q_{km} as a factor. Let $N_k{}^m$ be the $w_k \times w_k$ matrix that results from:

1) setting q_{km} equal to one in those elements of N_k in which it is a factor,
2) setting all other elements of N_k equal to zero.

For illustration let $l(k)=1$, $u(k)=5$ and $m=3$. Then, $j=3-1+1=3$. The fundamental matrix is

$$N_k(\Delta q_{km}) = \left[\begin{array}{ccc|cc} 1 & q_{k1} & q_{k1}q_{k2} & q_{k1}q_{k2}q_{k3} & q_{k1}q_{k2}q_{k3}q_{k4} \\ & 1 & q_{k2} & q_{k2}q_{k3} & q_{k2}q_{k3}q_{k4} \\ & & 1 & q_{k3} & q_{k3}q_{k4} \\ \cline{4-5} & & & 1 & q_{k4} \\ & & & & 1 \end{array}\right]$$

where $q_{k3} = \hat{q}_{k3} + \Delta q_{k3}$.

The elements having q_{k3} as a factor are in the indicated submatrix in the upper right corner of $K_k(\Delta q_{km})$. Setting q_{k3} equal to one in these elements and setting the other elements equal to zero we then have

$$N_k{}^m = \begin{bmatrix} 0 & 0 & 0 & q_{k1}q_{k2} & q_{k1}q_{k2}q_{k4} \\ 0 & 0 & 0 & q_{k2} & q_{k2}q_{k4} \\ 0 & 0 & 0 & 1 & q_{k4} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

From this illustration we observe that

$$N_k(\Delta q_{km}) = N_k + \Delta q_{km} N_k^{m} .$$

Note that N_k^{m} is not a function of q_{km}.

Thus, if $\Delta q_{lm}=\Delta p_{lm}=\Delta a_{lm}=0$ for $i=1,\cdots,n$ except that $\Delta q_{km}\neq 0$, then the stocks in grade i are from (8) and (9)

$$s_i(\Delta q_{km}) = f_1\left(\prod_{r=1}^{k-1} B_r\right)(N_k+\Delta q_{km} N_k^{m}) P_k \prod_{r=k+1}^{i-1} B_r N_i$$

$$= s_{i0} + \Delta q_{km} f_1\left(\prod_{r=1}^{k-1} B_r\right)(N_k^{m} P_k) \prod_{r=k+1}^{i-1} B_r N_i .$$

Define,

$$d_{ik} = f_1\left(\prod_{r=1}^{k-1} B_r\right)(N_k^{m} P_k) \prod_{r=k+1}^{i-1} B_r N_i \bar{1} , \text{ a scalar.} \tag{10}$$

Note that d_{ik} is not a function of q_{km} or of any Δq_{lm} or Δp_{lm}. So we have shown that when all Δq_{lm}'s and Δp_{lm}'s except Δq_{km} are zero then the total stocks in grade i are a linear function of Δq_{km}:

$$s_i(\Delta q_{km})\bar{1} = s_{i0} + d_{ik}\, \Delta q_{km} .$$

We now show the total stocks in grade i are also linear in Δp_{km} when all other Δq_{lm}'s and Δp_{lm}'s are zero.

The value of s_i depends on Δp_{km} only through $P_k(\Delta p_{km})$. Let P_k^{m} be the matrix that results from setting p_{km} equal to one in $P_k(\Delta p_{km})$ and setting all other elements of $P_k(\Delta p_{km})$ equal to zero. Then

$$P_k(\Delta p_{km}) = P_k(0) + \Delta p_{km} P_k^{m}$$

$$= P_k + \Delta p_{km} P_k^{m} .$$

Now if $\Delta q_{lm}=\Delta p_{lm}=\Delta a_{lm}=0$ for $l=1,\cdots,n$ except $p_{km}\neq 0$, then the stocks in grade i are

$$s_i(\Delta p_{km}) = f_1\Big(\sum_{r=1}^{k-1} B_r\Big)N_k(P_k+\Delta p_{km}P_k{}^m)\prod_{r=k+1}^{i-1} B_rN_i$$

$$= s_{i0} + \Delta p_{km}f_1\Big(\prod_{r=1}^{k-1} B_r\Big)N_kP_k{}^m\prod_{r=k+1}^{i-1} B_rN_i\ .$$

Define,

$$e_{ik} = f_1\Big(\prod_{r=1}^{k-1} B_r\Big)N_k\ P_k{}^m\prod_{r=k+1}^{i-1} B_rN_i\bar{1}\ . \qquad (11)$$

Note that e_{ik} is not a function of p_{km} or of any Δq_{lm} or Δp_{lm}.

So we have shown that when all Δq_{lm}'s and Δp_{lm}'s except Δp_{km} are zero then the total stocks in grade i are a linear function of Δp_{km}:

$$s_i(\Delta p_{km})\bar{1} = s_{i0} + e_{ik}\ \Delta p_{km}\ .$$

We call $N_k{}^m$ the m-differential matrix of N_k and $P_k{}^m$ the m-differential matrix of P_k.

The following lemma is used to show that when more than one Δq_{km} or Δp_{km} (or both) are non-zero, the total stocks are still linear functions of the Δq_{km}'s and Δp_{km}'s.

Lemma. If in the product $\prod_{k=1}^{i-1} B_kN_i$, where $B_k=N_kP_k$, any two (or more) distinct matrices are replaced by their m-differential matrices, then the product is a zero matrix.

Proof. Because all states have the 0-1 visiting property in the (grade, LOS) model, $\prod_{k=1}^{i-1} B_kN_i=V_{1i}$, the w_1 x w_i matrix

of probabilities of visiting a column state in grade i starting from a row state in grade 1. Each of these probabilities is the sum of the probabilities of all paths through the state space which start in the row state and end in the column state.

Replacing Q_k by q_{km} times its m-differential matrix Q_k^m is equivalent to extracting all terms in V_{1i} which have q_{km} as a factor. This in turn is equivalent to restricting the paths from the row state to the column state to those which pass through both states (k,m) and (k,m+1). Similarly, replacing P_k by p_{km} times its m-differential matrix P_k^m is equivalent to restricting the paths from the row state to the column state to those paths which pass through both states (k,m) and (k+1,m+1). Consequently, replacing two or more matrices in $\prod_{k=1}^{i-1} B_k N_i = V_{1i}$ by their m-differential matrices restricts the process to paths from the row state to the column state that have probability zero. $\square$

Thus we have.

$$s_i = f_1 \prod_{r=1}^{i-1} B_r(\Delta q_{rm}, \Delta p_{rm}) N_i(\Delta q_{im})$$

$$= f_1 \prod_{r=1}^{i-1} (N_r + \Delta q_{rm} N_r^m)(P_r + \Delta p_{rm} P_r^m)(N_i + \Delta q_{im} N_i^m).$$

From the lemma all of the "cross product" terms, i.e., terms containing two or more m-differential matrices, are zero matrices, so we have,

$$s_i = f_1 \prod_{r=1}^{i-1} N_r P_r N_i + \sum_{k=1}^{i} \Delta q_{km} f_1 \left(\prod_{r=1}^{k-1} B_r \right) N_k^m P_k \left(\prod_{r=k+1}^{i-1} B_r \right) N_i$$

$$+ \sum_{k=1}^{i-1} \Delta p_{km} f_1 \left(\prod_{r=1}^{k-1} B_r \right) N_k P_k^m \left(\prod_{r=k+1}^{i-1} B_r \right) N_i$$

and using (9), (10) and (11),

$$s_i \bar{1} = s_{io} \bar{1} + \sum_{k=1}^{i} \Delta q_{km} d_{ik} + \sum_{k=1}^{i-1} \Delta p_{km} e_{ik} .$$

This completes our demonstration that the constraints are linear in the decision variables. Thus the following program P2 is equivalent to the original program P1:

$$\text{P2]} \quad \min \sum_{k=1}^{n} c_k (\Delta q_{km} + \Delta p_{km})$$

$$\text{ST} \quad \sum_{k=1}^{i} \Delta q_{km} d_{ik} + \sum_{l=1}^{i-1} \Delta p_{km} e_{ik} \geq \bar{s}_i - s_{i0} \bar{1}, \quad i=1,\cdots,n ,$$

The constraints need not be of the exact form shown in this example. The nature of the problem does not change if the constraints are of the form

$$s_i L_i \geq \bar{s}_i , \quad i=1,\cdots,n$$

where L_i is a $w_i \times c_i$ matrix,

$\bar{s}_i$ is a $c_i \times 1$ vector,

and c_i is the number of constraints on stocks in grade i. The scalars d_{ik} and e_{ik} would in this case become $1 \times c_i$ vectors, and their defining equation would be modified by replacing $\bar{1}$ with L_i .

There are various efficient techniques for minimizing a nonlinear objective function subject to linear constraints which could be used to solve P2. The exact form of the

objective function might indicate the most appropriate technique, e.g., separable programming, quadratic programming or gradient projection. The point of the example, however, is to show that one may take advantage of the structure of the (grade, LOS) model to solve problems that would be quite difficult in a general two-characteristic model.

E. INTERCHANGING GRADE AND LOS

When studying manpower flows with a two-characteristic model it is sometimes advantageous to interchange the first characteristic, grade, and the secondary characteristic. The feasibility of this depends to a great extent on whether the interchanged characteristic model satisfies assumption A1. That is, in the original two-characteristic model, is the value of the second characteristic restricted to staying the same or increasing by one in each period?

In the case of the (grade, LOS) model interchanging grade and LOS does result in a model which satisfies assumption A1. We consider some of the properties of this interchanged model which we call the (LOS, grade) model.

Before interchanging characteristics it is convenient to expand the state space as necessary to insure that for each value of grade i and each value of LOS j there is a state (i,j). Recall that in the (grade, LOS) model we only define a state (i,j) when $l(i) \leq j \leq u(i)$. Let u be the largest of the u(i)'s. Then in the expanded state space (i,j) is defined for $1 \leq j \leq u$ for $i=1,\cdots,n$.

Now let us consider the structure of the submatrices in the overall transition matrix for the (LOS, grade) model.

For any LOS j the n x n matrix Q_j contains the transition probabilities for going from one grade to another while maintaining LOS constant at j. By the definition of LOS it must increase by one each period, so such probabilities are zero. Consequently, for $j=1,\cdots,u$,

$$Q_j = \bar{\bar{0}} ,$$

and,

$$N_j = (I-Q_j)^{-1} = I .$$

For any LOS j the n x n matrix P_j contains the probabilities of going from one pay grade to another (or the same pay grade) while increasing LOS to j+1. It is here that we can see a possible advantage of the (LOS, grade) model since we can allow demotions and multi-grade promotions without violating assumption A1. If there are no multi-grade promotions and there are no demotions, then P_j has non-zero entries only on and immediately above the main diagonal. If demotions are included in the model they cause non-zero entries below the main diagonal of P_j; promotions of any type are reflected in non-zero entries above the main diagonal of P_j. Under any circumstances,

$$B_j = N_j P_j = P_j .$$

If demotions are not included in the model, then B_j is upper triangular and so is the product of successively indexed matrices B_j.

For any LOS j the n x 1 matrix A_j contains the probabilities of leaving the system from the various pay grades when length of service is j.

The overall transition matrix P for the (LOS, grade) model has the same form as in any secondary characteristic model. The important difference is that in the (LOS, grade) model we have $Q_j=\bar{\bar{0}}$. Thus,

$$P = \left[\begin{array}{ccccc|c}
\bar{\bar{0}} & P_1 & & & & A_1 \\
 & \bar{\bar{0}} & P_2 & & & A_2 \\
 & & \ddots & \ddots & & \vdots \\
 & & & \ddots & P_{u-1} & A_{u-1} \\
 & & & & \bar{\bar{0}} & A_u \\
\hline
 & \bar{0} & & & & 1
\end{array}\right].$$

To illustrate the effects of interchanging grade and LOS consider an example from the U.S. Marine Corps. The enlisted force of the Marine Corps has nine grades. The values of LOS range from 1 to 30. So in the (grade, LOS) model we would have nine 30x30 matrices Q_i and eight 30x30 matrices P_i. Of course each of these matrices would have at most 30 non-zero elements. By interchanging characteristics, the (LOS, grade) model has all matrices Q_i equal to zero matrices, and there are twenty-nine 9x9 matrices P_i. Whether it is more practical to use nine pairs of relatively large matrices or twenty-nine relatively small matrices must depend on the problem to be solved. This example does illustrate that the modeler has a choice.

The fundamental matrix for the transient part of the process is,

$$N = \left[\begin{array}{ccccc} I & P_1 & P_1P_2 & --- & \prod_{j=1}^{u} P_j \\ & I & P_2 & --- & \prod_{j=2}^{u} P_j \\ & & I & --- & \prod_{j=3}^{u} P_j \\ & & & \ddots & \vdots \\ & & & & I \end{array}\right].$$

The expansion of the state space so that all states in $\{(i,j):j=1,\cdots,u\ ,\ \ i=1,\cdots,n\}$ are defined is not always necessary. In the (LOS, grade) model one may define,

$l(j)$ = lowest grade that a person having LOS j may hold,

$u(j)$ = highest grade that a person having LOS j may hold.

As before $w_j=u(j)-l(j)+1$. Then Q_j is $w_j \times w_j$, P_j is $w_j \times w_{j+1}$ and A_j is $w_j \times 1$. The nature of the model is not substantially affected. Computer storage requirements may be reduced but computer programming may be more complex.

We note that the basic stock equation (see Chapter III, equation (1)) in the (LOS, grade) model is

$$s_j(t+1) = f_j(t+1) + s_{j-1}(t)P_{j-1}\ ,$$

reflecting the fact that the stocks of LOS j in the next period do not depend on the stocks of LOS j in the present period. The cumulative stock equation (equation (2), Chapter III) is the same as the basic stock equation. The following equation is derived from the fundamental matrix N_3 and gives

the stocks in terms of the external flows in previous periods,

$$s_j(t) = \sum_{m=1}^{j} f_m(t-m+j) \prod_{k=m}^{j-1} P_k .$$

In summary the choice between the (grade, LOS) or the (LOS, grade) model depends on the information that is to be derived from the model. The (LOS, grade) model is more flexible in that the promotion/demotion structure is not restricted. The (grade, LOS) model might be more tractable when used as part of a budget model since costs are usually dependent on grade rather than on LOS. Under any circumstances the two models must yield equivalent information.

V. THE (GRADE, TIG) MODEL

The (grade, TIG) model is a model of a graded manpower system in which the second characteristic is time in grade (TIG). By time in grade we mean the number of periods that a person has been in his present grade.

A. BACKGROUND

There are occasions when it is desired to model the flows of "career motivated" people in a manpower system. The operational definition of a career motivated person depends on the case at hand, but generally we expect such people to have no predilection to leave the system at the earliest convenient opportunity. Rather we expect the decisions of career motivated people with respect to leaving the system to be closely related to the recognition they receive for their performance in the system. One measure of recognition of performance is the combination of pay grade and time since last promotion. Such considerations may make a (grade, TIG) model appropriate.

A somewhat different context in which the (grade, TIG) model may be appropriate is that in which the organization has an "up or out"policy. Positions in an organization may be (formally or informally) partitioned between those that are on the "path towards the top" and those that are terminal positions. The former group often has an up or out policy where "out" means transfer to a terminal position; the

(grade, TIG) model can be used to examine flows in positions in this group.

B. DESCRIPTION

A person who enters grade i during period t is counted for the first time in grade i at the end of the period (at time t) and assigned a TIG of one at that time. The person's TIG increases by one for each successive period that he is counted in grade i.

The states of the system are defined by couples (i,j) where:

(i,j) = the state corresponding to grade i and TIG j.

The notation and results for the two-characteristic model apply directly to the (grade, TIG) model. In the (grade, TIG) model. In the (grade, TIG) model the value of l(i) is always one.

We need define only the following probabilities:

q_{ij} = probability a person in state (i,j) at the end of one period wll be in state (i,j+1) at the end of the next period.

p_{ij} = probability a person in state (i,j) at the end of one period will be in state (i+1,1) at the end of the next period.

a_{ij} = probability a person in state (i,j) at the end of one period will be out of the system at the end of the next period.

Under assumption A1,

$$q_{ij} + p_{ij} + a_{ij} = 1.$$

The transition matrix A_i is a $w_i \times 1$ matrix:

$$A_i = [a_{i1}, \cdots, a_{i,u(i)}] .$$

The transition matrix Q_i is $w_i \times w_i$ and has non-zero elements only immediately above the main diagonal:

$$Q_i = \begin{bmatrix} 0 & q_{i1} & & & \\ & 0 & q_{i2} & & \\ & & 0 & \ddots & \\ & & & \ddots & q_{i,u(i)-1} \\ & & & & 0 \end{bmatrix} .$$

The transition matrix P_i is $w_i \times w_i$ and has non-zero elements only in the first column:

$$P_i = \begin{bmatrix} p_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ p_{i,u(i)} & 0 & \cdots & 0 \end{bmatrix} .$$

Thus, A_i and Q_i have the same form as in the (grade, LOS) model. Each fundamental matris N_i also has the same form as in the (grade, LOS) model:

$$N_i = \begin{bmatrix} 1 & q_{i1} & q_{i1}q12 & \cdots & \prod_{j=1}^{u(i)-1} q_{ij} \\ & 1 & q_{i2} & \cdots & \prod_{j=2}^{u(i)-1} q_{ij} \\ & & \ddots & & \vdots \\ & & & \ddots & 1 \end{bmatrix} .$$

The matrix $B_i=N_iP_i$ has non-zero elements only in the first column because grade i+1 can only be entered in state (i+1,1). Let the element in row j and column 1 of B_i be denoted b_{ij}. Then,

$$b_{ij} = \sum_{k=j}^{u(i)} \left(\prod_{m=j}^{k-1} q_{im} \right) p_{ik} .$$

Recall that $B_i\bar{1}$ is a $w_i \times 1$ vector, and its j^{th} component is the probability of attaining grade i+1 starting from state (i,j). In the (grade, TIG) model,

$$B_i\bar{1} = [b_{i1},\cdots,b_{i,u(i)}] ,$$

so,

b_{ij} = probability of attaining grade i+1 from state (i,j).

The product of successively indexed B_i matrices has non-zero elements only in the first column, and

$$\prod_{m=i}^{k-1} B_m = \begin{bmatrix} b_{i1} \prod_{m=i+1}^{k-1} b_{m1} & 0 - - - 0 \\ b_{i2} \prod_{m=i+1}^{k-1} b_{m1} & 0 - - - 0 \\ \vdots & \vdots \quad\quad \vdots \\ b_{i,u(i)} \prod_{m=i+1}^{k-1} b_{m1} & 0 - - - 0 \end{bmatrix}$$

$$= \prod_{m=i+1}^{k-1} b_{m1} [B_i\bar{1} , \bar{\bar{0}}] .$$

If $q_{ij}=0$, then state (i,j+1) is unreachable (except possibly through an external flow). If $q_{ij}\neq 0$, then,

$$b_{i,j+1} = \frac{b_{ij}-p_{ij}}{q_{ij}} .$$

This relation provides an efficient method for computing the elements of the vector $B_i\bar{1}$ recursively:

$$b_{i,u(i)} = p_{i,u(i)} ,$$

$$b_{ij} = p_{ij} + q_{ij}b_{i,j+1} , \quad j=u(i)-1,\cdots,1 .$$

The (grade, TIG) model is equivalent to a discrete semi-Markov process. The states of the process are the n grades and T_0 (out of the system). Upon entering state i the next transition is to state i+1 with probability b_{i1} and to T_0 with probability $1-b_{i1}$. The tail distribution of the time spent in grade i is the first row of the fundamental matrix N_i. The results from the two-characteristic model in Chapter II, Section E can be used to find the distributions of time spent in state i conditioned on either promotion or leaving the system before promotion. Thus all the information required to set up the (grade, TIG) model as a discrete semi-Markov process is readily derived from the results established here. The use of the discrete semi-Markov process is covered in some detail in Howard, 1971. Because of assumption A1 there seems to be little advantage to treating the (grade, TIG) model as a semi-Markov process. If it is necessary that demotions and multi-step promotions be included in a model based on grade and time in grade, then the more general techniques of semi-Markov processes might be of value. We show in the following section, however, that interchanging characteristics in the (grade, TIG) model leads to a Markov model which allows the inclusion of demotions and multi-step promotions. This model is

computationally practical, but violates assumption A1, and thus is not in the class of models analyzed in Chapters II and III.

C. INTERCHANGING GRADE AND TIG

Interchanging grade and time in grade results in a model that does not satisfy assumption A1 (the new first characteristic, time in grade, decreases when the new second characteristic, grade, changes). Nevertheless, the (TIG, grade) model is not too difficult to analyze, and it enables the modeler to provide for demotions and multi-grade promotions.

Let (j,i) denote the state in the (TIG, grade) model corresponding to j periods in grade i. As before i is in the set $\{1,\cdots,n\}$, and we will let j take any value from the set $\{1,\cdots,m\}$ where the value of m is specified. Let P_j be an $n \times n$ matrix of transition probabilities for transitions from states in $\{(j,i):i=1,\cdots,n\}$ to states in $\{(1,k):k=1,\cdots,n\}$. That is, P_j is the transition matrix for transitions from one grade to a different grade when TIG is equal to j. Changing grades, of course, causes the value of TIG to change to one. The matrix P_j has zeroes on its main diagonal. Let Q_j be an $n \times n$ matrix of transition probabilities for transitions from states in $\{(j,i):i=1,\cdots,n\}$ to states in $\{(j+1,i):i=1,\cdots,n\}$. That is, Q_j represents transitions in which grade is unchanged and TIG increases from j to j+1. The matrix Q_j has zeroes everywhere except possibly on its main diagonal. Let A_j be an $n \times 1$ matrix of transition probabilities for transitions from states in $\{(j,i):i=1,\cdots,n\}$ to out of the system. The

overall transition matrix P for the (TIG, grade) model is shown below:

$$P = \left[\begin{array}{ccccc|c}
P_1 & Q_1 & & & & A_1 \\
P_2 & & Q_2 & & & A_2 \\
\vdots & & & \ddots & & \vdots \\
P_{m-1} & & & & Q_{m-1} & A_{m-1} \\
P_m & & & & & A_m \\
\hline
 & & 0 & & & 1
\end{array}\right].$$

The transient part of the process is represented by the submatrices in the upper-left part of P. We call this (as before) the transient matrix Q. By assumption A2 the process is transient, so the fundamental matrix, $N=(I-Q)^{-1}$, exists. Of particular interest is the $n \times n$ submatrix in the upper-left corner of N which we will denote by N_{11}. The rows and columns of N_{11} correspond to states in $\{(1,i):i=1,\cdots,n\}$. The basic theorem on the fundamental matrix [Theorem 3.2.4, Kemeny and Snell, 1960] indicates that the element in row i and column k of N_{11} equals the expected number of visits to state (1,k) starting from state (1,j). If the 0-1 visiting property holds (which in the (TIG, grade) model means no demotions), then the element in row i and column k of N_{11} equals the probability of attaining grade k given that grade i has just been attained. The submatrix N_{11} may be computed in the following manner.

Let,

$F_m = P_m$,

$F_j = P_j + Q_j F_{j+1}$, $j=1,\cdots,m-1$, an $n \times n$ matrix.

The element in row i and column k of F_j is the probability of ever entering state (1,k) starting from state (j,i). It can be shown that,

$$N_{11} = (I-F_1)^{-1},$$

and that the first column of submatrices of N, an $mn \times n$ matrix, is $[N_{11}, F_2N_{11}, \cdots, F_mN_{11}]$. Most of the first-order information of interest from the (TIG, grade) model is either contained in or readily derived from the first column of submatrices of the fundamental matrix N.

In summary, interchanging characteristics in the (grade, TIG) model leads to a model having a transient matrix structure quite different from that of the two-characteristic model as developed in Chapter II. However, first-order properties of the interchanged model are not too difficult to obtain. The most obvious advantage of interchanging characteristics is that demotions and multi-grade promotions may be included in the model.

VI. COMBINING STATES

In this chapter we consider some of the mathematical properties of the two-characteristic model when the states in each grade are combined into a single state. The ideas discussed here are closely related to the concept of lumpability as presented in Burke and Rosenblatt, 1958, and Kemeny and Snell, 1960. There is a difference between the approach taken here and that of the foregoing authors. Burke, et al. are primarily interested in establishing conditions under which the combining of states in a stationary Markov process leads to a process which is still Markov and stationary and has these properties for all (or at least some) initial probability vectors. In Section A of this chapter we briefly consider conditions under which the states of a two-characteristic model are lumpable. In Section B of this chapter we will consider combining states with the clear understanding that the resulting process may not be Markov or stationary for most initial probability vectors. Section C contains an example which illustrates the results in Section B.

We restrict our attention to the combining of all states in each grade so that the "reduced state space" version of the two-characteristic model will have one state for each grade and a state for "out of the system." Obviously such a combining of states leads to a one-characteristic model.

A. LUMPABILITY

Let $\phi(m)$ be the random function which indicates the state that a Markov process is in after m steps. Recall that T_i is the set of all states in grade i. Let π be an initial probability vector over all states in the system. Then let

(1) $\Pr[\phi(0) \varepsilon T_i]$ = probability the process starts in T_i given the initial probability vector π,

(2) $\Pr_\pi[\phi(m+1) \varepsilon T_t | \phi(m) \varepsilon T_s, \cdots, \phi(1) \varepsilon T_r, \phi(0) \varepsilon T_i]$
= probability the process is in T_t after m+1 steps, given the initial probability vector π and the events $\phi(0) \varepsilon T_i$, $\phi(1) \varepsilon T_r, \cdots, \phi(m) \varepsilon T_s$.

The latter probability is not defined unless the given sequence of events has positive probability under the initial probability vector π.

Let T* denote the partition of the state space $T \cup T_0$ into the sets of states, $T_0, T_i, \cdots, T_n$. The foregoing probabilities (1) and (2) define a stochastic process on T*, and we call this a lumped process.

Definition. We say that the Markov chain in a two-characteristic model is "lumpable" with respect to the partition T* if for every starting vector π the lumped process defined above is a Markov chain and the transition probabilities do not depend on the choice of π.

Theorem. A necessary and sufficient condition for the Markov chain in a two-characteristic model to be lumpable with respect

to T* is that for every pair of sets in T*, e.g., T_i and T_k, the probability of a one-step transition from any state $(i,j) \epsilon T_i$ to some state in T_k has the same value for every state in T_i. These common values, denoted by $\{\hat{p}_{ik}\}$, form a transition matrix for the lumped chain.

The foregoing definition and theorem are taken from Chapter VI of Kemeny and Snell, 1960, with paraphrasing to suit the case at hand.

By assumption A1 for any starting vector ,

$$\Pr_{\pi}[\phi(m+1) \epsilon T_t \,|\, \phi(m) \epsilon T_s, \cdots, \phi(1) \epsilon T_r,\ \phi(0) \epsilon T_i]$$

can be non-zero only if t=s,s+1, or 0.

We define the w_i x 3 matrix $\hat{P}_i$ by,

$$\hat{P}_i = [Q_i\bar{1},\ P_i\bar{1},\ A_i\bar{1}]. \tag{1}$$

Corollary 1. The Markov chain in a two-characteristic model is lumpable with respect to T* if and only if for each grade $i=1,\cdots,n$ the matrix $\hat{P}_i$ has identical rows. If the chain is lumpable, then the transition matrix for the chain is

$$\hat{P} = \left[\begin{array}{cccccc|c}
\hat{p}_{11} & \hat{p}_{12} & & & & & \hat{p}_{10} \\
 & \hat{p}_{22} & \hat{p}_{23} & & & & \hat{p}_{20} \\
 & & \ddots & \ddots & & & \vdots \\
 & & & & & \hat{p}_{n-1,n} & \hat{p}_{n-1,0} \\
 & & & & & \hat{p}_{n,n} & \hat{p}_{n,0} \\
\hline
 & & \bar{0} & & & & 1
\end{array}\right]$$

where the vector $(\hat{p}_{ii}, \hat{p}_{i,i+1}, \hat{p}_{i0})$ is any one of the identical rows of $\hat{p}_i$.

Corollary 2. A necessary condition for the Markov chain in a (grade, LOS) model or a (grade, TIG) model to be lumpable is that $Q_i=\bar{\bar{0}}$, $i=1,\cdots,n$.

Proof. In the (grade, LOS) and (grade, TIG) models $q_{i,u(i)}=0$ by definition. So,

1) the last component of $Q_i\bar{1}$ is a zero, implying
2) the rows of $\hat{P}_i$ are not identical unless $q_i\bar{1}=0$, implying,
3) the rows of $\hat{P}_i$ are not identical unless $Q_i=\bar{\bar{0}}$, implying by Corollary 1,
4) the chain is not lumpable with respect to T^* unless $Q_i=\bar{\bar{0}}$, $i=1,\cdots,n$. $\Box$

In the (grade, LOS) and (grade, TIG) models if $Q_i=\bar{\bar{0}}$, $i=1,\cdots,n$, the lumped process is rather trivial because the only sequences of events that have positive probability are those of the form $\phi(0)\varepsilon T_i,\phi(1)\varepsilon T_{i+1},\cdots,\phi(m)\varepsilon T_{i+m}$, or such a sequence with events $\phi(m+k)\varepsilon T_0$, $k=1,\cdots,r$ appended. If the starting vector π has positive probability only on T_i or we are given that $\phi(0)\varepsilon T_i$, then the lumped process is Markovian but the transition probabilities depend on the value of π, a situation Kemeny and Snell call "weak lumpability."

B. LUMPING AN UNLUMPABLE PROCESS

It may be necessary or advantageous in some circumstances to use a one-characteristic model. If the system is more accurately modeled by a particular two-characteristic model, then we are faced with the problem of lumping a process which may be mathematically unlumpable. In this section we develop

methods for qualitatively judging what parameters might be appropriate when lumping a two-characteristic model. It is convenient to abandon the stochastic interpretation of the model, and adopt a deterministic fractional flow viewpoint of the process.

Recall equation (1), $\hat{P}_i = [Q_i\bar{1}, P_i\bar{1}, A_i\bar{1}]$, a $w_i \times 3$ matrix. In this section we assume that the lumpability condition is not met, i.e., the rows of $\hat{P}_i$ are not identical. We are interested in the following fractions for each grade $i=1,\cdots,n$:

1) the fraction of those in grade i at t who remain in grade i for one more period,

2) the fraction of those in grade i at t who get promoted to grade i+1 during period t+1,

3) the fraction of those in grade i at t who leave the system during period t+1.

The values of these fractions depend on the distribution of the stocks in grade i at time t.

For any stock vector $s_i(t) \neq \bar{0}$, we define,

$$\tilde{s}_i(t) = \frac{s_i(t)}{s_i(t)\bar{1}}, \text{ a } 1 \times w_i \text{ distribution vector.} \tag{2}$$

Then,

$$\tilde{s}_i(t)P_i = \left(\frac{s_i(t)Q_i\bar{1}}{s_i(t)\bar{1}}, \frac{s_i(t)P_i\bar{1}}{s_i(t)\bar{1}}, \frac{s_i(t)A_i\bar{1}}{s_i(t)\bar{1}}\right), \text{ a } 1 \times 3 \text{ vector.}$$

Define the following functions which map distribution vectors into scalars:

$$q_i(\tilde{s}_i(t)) = s_i(t)Q_i\bar{1},$$

$$p_i(\tilde{s}_i(t)) = s_i(t)P_i\bar{1},$$

$$a_i(\tilde{s}_i(t)) = s_i(t)A_i\bar{1}.$$

If the vector of stocks in grade i at time t is $\tilde{s}_i(t) \neq 0$ and $\tilde{s}_i(t)$ is defined by equation (2), then,

$q_i(\tilde{s}_i(t))$ = the fraction of those in grade i at t who remain in grade i for one more period,

$p_i(\tilde{s}_i(t))$ = the fraction of those in grade i at t who get promoted to grade i+1 during period t+1,

$a_i(\tilde{s}_i(t))$ = the fraction of those in grade i at t who leave the system during period t+1.

We note that,

$$q_i(\tilde{s}_i(t))+p_i(\tilde{s}_i(t))+a_i(\tilde{s}_i(t)) = 1.$$

Define the following function which maps a $1 \times w_i$ distribution vector into a 1×3 distribution vector:

$$\begin{aligned}\hat{P}_i(\tilde{s}_i(t)) &= \tilde{s}_i(t)\hat{P}_i \\ &= (q_i(\tilde{s}_i(t)),\ p_i(\tilde{s}_i(t)),\ a_i(\tilde{s}_i(t))). \qquad (3)\end{aligned}$$

If the stock distribution vectors for grade i, $\tilde{s}_i(t)$, were known for $t=0,1,2,\cdots$, then one could combine the states in grade i and use the sequence $\langle\hat{p}_i(\tilde{s}_i(t)\rangle_{t=0}^{\infty}$ to form a non-stationary fractional flow model.

The more pertinent situation is that in which $s_i(0)$ is known, but we do not know or do not want to compute $s_i(t)$ and $\tilde{s}_i(t)$ for $t \geq 1$. Then in the absence of information about future stock vectors we want to choose the parameters of a stationary lumped process in some reasonable fashion. That

is, for each grade $i=1,\cdots,n$ we want to select three numbers (which we will denote q_i, p_i and a_i) to represent the fractions of the people in grade i who stay in grade i, get promoted and leave the system respectively each period. The lumped process would have a state for each grade and one state for out of the system. The transition matrix would be,

$$\hat{P} = \left[\begin{array}{ccccc|c}
q_1 & p_1 & & & & a_1 \\
 & q_2 & p_2 & & & a_2 \\
 & & \ddots & \ddots & & \vdots \\
 & & & \ddots & p_{n-1} & a_{n-1} \\
 & & & & q_n & a_n \\ \hline
 & \bar{0} & & & & 1
\end{array}\right].$$

Under the assumption that the lumpability condition is not met, there is no choice of parameters for $\hat{P}$ that will in general be equivalent to the unlumped process. We may begin, however, by eliminating choices for $\hat{P}$ that are "obviously bad."

For any distribution vector $\tilde{s}_i(t)$, the 1 x 3 vector $\hat{p}_i(\tilde{s}_i(t))$ is a convex combination of the rows of $\hat{p}_i$. Let H_i be the set of all such vectors, i.e.,

$$H_i = \{h_i \varepsilon E^3 : h_i = \tilde{s}_i\hat{P}_i,\ \tilde{s}_i \quad 0,\ \tilde{s}_i\bar{1} = 1\}.$$

In choosing parameters for the lumped process, any choice (q_i, p_i, a_i) which is not in H_i is "obviously bad" because there is no stock vector for which this choice reflects the behavior of the system.

The set H_i is a subset of the fundamental simplex in 3-space.

Clearly (from Corollary 2) the Markov chain of a two-characteristic model is lumpable if and only if for each grade $i=1,\cdots,n$, H_i is a single point. A plot of H_i gives some qualitative indication of just how unlumpable grade i actually is.

Next we consider the behavior of the sequence $<\hat{P}_i(\tilde{s}_i(t))>$ under various assumptions.

1. Constant Stocks

If the stock vector does not change with time, i.e., $s_i(t) = s_i$, $t=0,1,2,\cdots$, then the stock distribution vector is constant,

$$\tilde{s}_i(t) = \frac{s_i}{s_i\bar{1}} \equiv \tilde{s}_i ,$$

and the logical choice of parameters for the lumped process is,

$$\begin{aligned}(q_i, p_i, a_i) &= \hat{P}_i(\tilde{s}_i) \\ &= \frac{s_i}{s_i\bar{1}} \hat{P}_i ,\end{aligned}$$

because these parameters exactly reflect the deterministic flow behavior of the system.

2. Convergent Stock Distribution Vector

If the sequence of stock vectors, $<s_i(t)>$, has the property that the corresponding sequence of stock distribution vectors, $<\tilde{s}_i(t)>$, converges to a distribution vector $\tilde{s}_i$, then the sequence of linear transformations of $\tilde{s}_i(t)$, $<\hat{P}_i(\tilde{s}_i(t))>$, converges to $\hat{P}_i(\tilde{s}_i)$. In this case it may be reasonable to

choose the parameters of the lumped process by taking a convex combination of $\hat{P}_i(\tilde{s}_i(0))$ and $\hat{P}_i(\tilde{s}_i)$. For example,

$$(q_i,\ p_i,\ a_i) = (1-\theta)\hat{P}_i(\tilde{s}_i(0)) + \theta\hat{P}_i(\tilde{s}_i),\ 0 \leq \theta \leq 1,$$

where small values of θ are appropriate for short-range planning and values of θ close to 1 are more appropriate for long-range planning.

3. Constant External Flows

It was shown in Chapter III that under constant external flows, i.e.,

$$f_i(t) = f_i, \text{ a } 1 \times w_i \text{ nonnegative vector,}$$

the limiting value of the stock vector is

$$s_i = \left(\sum_{k=1}^{i} f_k B_{ki}\right)N_i .$$

We define

$$g_i = \sum_{k=1}^{i} f_k B_{ki} , \tag{4}$$

a $1 \times w_i$ vector of inputs into each state in grade i in steady state. We assume $g_i \neq \bar{0}$. Then,

$$\begin{aligned} \tilde{s}_i &= s_i/s_i\bar{1} \\ &= g_iN_i/(g_iN_i\bar{1}) \\ &= \left(\frac{g_i}{g_i\bar{1}}\ N\right)/\left(\frac{g_i}{g_i\bar{1}}\ N_i\bar{1}\right. , \end{aligned} \tag{5}$$

so with respect to $\tilde{s}_i$ we may without loss of generally assume that g_i is a distribution vector.

Let

$$H_i^* = \{h_i \varepsilon E^3 :\ h_i^* = \tilde{s}_i\hat{P}_i,\ \tilde{s}_i \text{ calculated from (4) and (5)}\}.$$

That is, H_i^* is the set of limit points of the sequence $\langle \hat{p}_i(\tilde{s}_i(t)) \rangle$ under constant external flows. We show that (except for a special case) the set H_i^* is a proper subset of H_i.

Let,

N_{ij} = j^{th} row of N_i, a $1 \times w_i$ vector,

g_{ij} = j^{th} element of g_i, a scalar,

$\tau_{ij} = N_{ij}\bar{1}$, a scalar. (See equation (4) of Chapter II.) From Chapter II we have that τ_{ij} is the average time spent in grade i starting from state $(i, j+l(i) - 1)$. Note that $\tau_{ij} \geq 1$. Let D_i be a $w_i \times w_i$ matrix having $\{\tau_{ij}; j=1,\cdots,w_i\}$ on its main diagonal and zeroes elsewhere. Then D_i^{-1} is defined.

For any $1 \times w_i$ nonnegative vector g_i, let

$$g_i^* = \frac{1}{g_i N_i \bar{1}} (g_{i1}\tau_{i1}, \cdots, g_{i,w_i}\tau_{i,w_i}), \text{ a } 1 \times w_i \text{ vector.} \qquad (6)$$

Then it can be shown that g_i^* is a distribution vector, and,

$$\begin{aligned}\tilde{s}_i &= g_i N_i/(g_i N_i \bar{1}) \\ &= g_i^* D_i^{-1} N_i . \qquad (7)\end{aligned}$$

Let,

$S(w_i)$ = {all $1 \times w_i$ distribution vectors},

$\tilde{S}(w_i)$ = {all $1 \times w_i$ distribution vectors
determined by (7)}.

Clearly, $\tilde{S}(w_i)$ is a subset of $S(w_i)$; we now establish conditions under which it is a proper subset.

Equation (6) maps $S(w_i)$ into $S(w_i)$, and it can be shown that the mapping is onto. That is, for every g_i^* in $S(w_i)$ there is a g_i in $S(w_i)$ such that equation (6) maps g_i to g_i^*. Thus, equation (7) may be viewed as a mapping from $S(w_i)$ into $\tilde{S}(w_i)$. We show, however, that the mapping in equation (7) may not be onto; i.e., there may exist distribution vectors in $S(w_i)$ which are not limiting stock distributions for any choice of constant external flows (recall that (4) gives g_i as a function of the constant external flows).

The $w_i \times w_i$ matrix $(D_i^{-1}N_i)$ is a non-singular linear transformation from $S(w_i)$ into $\tilde{S}(w_i)$. It is well-known (see, for example, Halmos, 1956) that for any non-singular linear transformation T having determinant d and any measurable set E, with Lebesque measure m(E), $m(E)/m(T^{-1}E) = |d|$. In the present context this can be written,

$$\frac{m(\tilde{S}_i(w_i))}{m(S_i(w_i))} = |\det(D_i^{-1}N_i)|.$$

But the matrix $(D_i^{-1}\,N_i)$ is nonnegative and its rows all sum to one (i.e., it is a stochastic matrix). So its eigenvalues are bounded in absolute value by +1. The determinant of a matrix equals the product of its eigenvalues, so,

$$\frac{m(\tilde{S}_i(w_i))}{m(S_i(w_i))} = |\det(D_i^{-1}N_i)| \leq 1.$$

Furthermore, $m(\tilde{S}_i(w_i)) = m(S_i(w_i))$ if and only if the magnitude of every eigenvalue of $(D_i^{-1}\,N_i)$ is 1. It can be shown (see Karlin, theorem 2.1, page 97), that this can only occur when

the non-zero elements of Q_i are restricted to the main diagonal. This would mean that a person entering grade i in state (i,j) could not make a transition to any other state in grade i.

If $|\det(D_i^{-1} N_i)|<1$, then $\tilde{S}_i(w_i)$ is a proper subset of $S_i(w_i)$. The sets H_i and H_i^* are linear transformations of $S(w_i)$ and $\tilde{S}(w_i)$ respectively, so $|\det(D_i^{-1} N_i)|<1$ implies that H_i^* is a proper subset of H_i.

When the set of states in grade i has the 0-1 visiting property we may assume without loss of generality that N_i is upper triangular with ones on its main diagonal (see Chapter II, Section C). Then the determinant of N_i is +1. The determinant of the diagonal matrix D_i is,

$$\det D_i = \prod_{j=1}^{w_i} \tau_{ij} .$$

But $\tau_{ij} \geq 1$, so $\det D_i = 1$ if and only if all τ_{ij}'s equal 1. When the 0-1 visiting property holds this requires that Q_i be a zero matrix. Thus, if the 0-1 visiting property holds and Q_i is not a zero matrix, then

$$\frac{m(\tilde{S}_i(w_i))}{m(S_i(w_i)} = \left(\prod_{j=1}^{w_i} \tau_{ij}\right)^{-1} < 1.$$

Furthermore any change in Q_i that increases the average time spent in grade i, $\{\tau_{ij}\}$, will decrease the size of $\tilde{S}_i(w_i)$ and thus decrease the size of H_i^*.

In summary, it has been shown that under constant external flows the vector of the fractions retained in grade,

promoted and lost from the system converges to some point in H_i^* which is in most cases of interest a proper subset of H_i. Thus, when the lumped process is to be used for long-range projections, choosing the parameter vector (q_i, p_i, a_i) from H_i^* would seem quite reasonable.

4. Linear Growth of External Flows

The case of linear growth of the external flows leads to results similar to those in the case of constant external flows. It was shown in Section C of Chapter III that under linear growth of external flows, $f_1(t) = tf_i$, the stock vector in grade i is asymptotically linear and

$$s_i(t) = tL_i + C_i$$

where

$$L_i = (f_i + L_{i-1}P_{i-1})N_i,$$

$$C_i = -((L_{i-1} - C_{i-1})P_{i-1} + f_iN_iQ_i)N_i \; .$$

The limiting stock distribution vector is

$$\begin{aligned} \lim_{t\to\infty} \tilde{s}_i(t) &= \lim_{t\to\infty} s_i(t)/s_i(t)\bar{1} \\ &= \lim_{t\to\infty} (tL_i + C_i)/(tL_i + C_i)\bar{1} \\ &= L_i/L_i\bar{1} \; . \end{aligned}$$

But,

$$L_i = \sum_{k=1}^{i} f_k \, B_{ki} \, N_i \; .$$

So the limiting value of $\tilde{s}_i(t)$ is the same in the case of linear-growth external flows as it was in the case of constant external flows. Consequently the set of all possible

limit points for the flow fractions, H_i^*, is the same under both constant and linear-growth external flows.

C. AN EXAMPLE

In this section we use an example to illustrate how the planner might use the results of Section B to combine states in a two-characteristic model.

Let,

$$Q_i = \begin{bmatrix} 0 & .2 & 0 & 0 \\ 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & .4 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad P_i = \begin{bmatrix} .7 & 0 & 0 & 0 \\ 0 & .4 & 0 & 0 \\ 0 & 0 & .0 & 0 \\ 0 & 0 & 0 & .1 \end{bmatrix}, \quad A_i = \begin{bmatrix} .1 \\ .1 \\ .5 \\ .9 \end{bmatrix}.$$

Then,

$$P_i = [Q_i\bar{1},\ P_i\bar{1},\ A_i\bar{1}]$$

$$= \begin{bmatrix} .2 & .7 & .1 \\ .5 & .4 & .1 \\ .4 & .1 & .5 \\ 0 & .1 & .9 \end{bmatrix}.$$

Clearly, the rows of $\hat{P}_i$ are not identical, so the condition for lumpability is not satisfied. The rows of $\hat{P}_i$ are the extreme points of H_i, ahd H_i is plotted in solid lines in Figure 2. Any choice of parameters for the combined process should be taken from this set H_i.

Let the initial stock vector be,

$$s_i(0) = (20,\ 70,\ 10,\ 0).$$

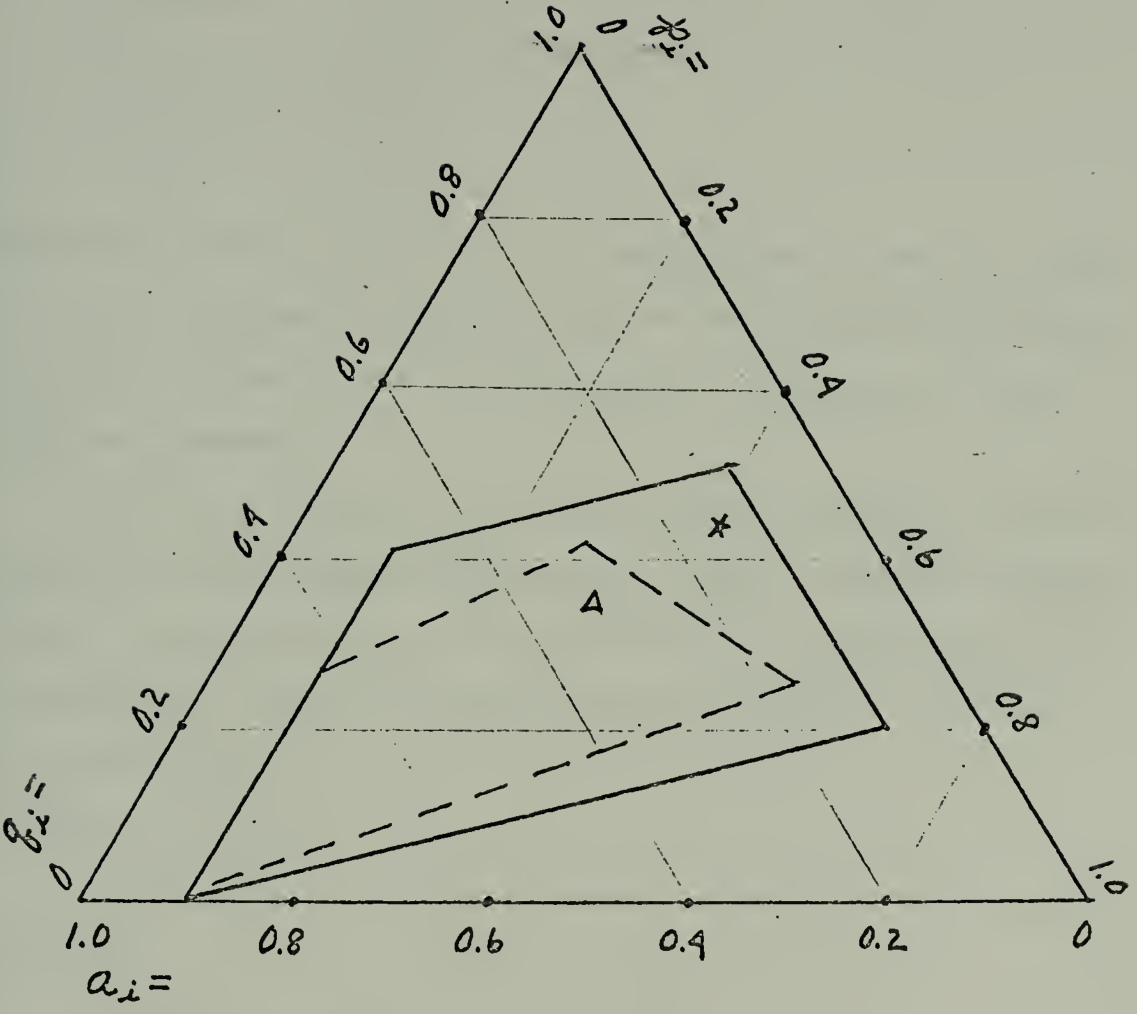

Figure 2. Plot of Points and Sets Used in

in Combining States.

Then,

$$\tilde{s}_i(0) = s_i(0)/s_i(0)\bar{1}$$

$$= (0.2,\ 0.7,\ 0.1,\ 0),$$

and,

$$\tilde{s}_i(0)\ \hat{P}_i = (0.43,\ 0.43,\ 0.14).$$

That is,

$$q_i\ (\tilde{s}_i(0)) = 0.43,$$

$$p_i\ (\tilde{s}_i(0)) = 0.43,$$

$$a_i\ (\tilde{s}_i(0)) = 0.14.$$

The point (0.43, 0.43, 0.14) is denoted by a star in Figure 2. For a short-range projection (one or two periods) this point would be a good choice for the grade i parameters of the combined process.

Next let us consider what choices of parameters would be appropriate for long-range planning under constant external flows. From equation (7) of Section B.2 we know that the distribution vector for the equilibrium stocks in grade i, $\tilde{S}_i$, must satisfy,

$$\tilde{s}_i = g_i^* D_i^{-1} N_i ,$$

where g_i^* is a distribution vector. In the present example,

$$N_i = \begin{bmatrix} 1 & 0.2 & 0.1 & 0.04 \\ 0 & 1 & 0.5 & 0.2 \\ 0 & 0 & 1 & 0.4 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and the diagonal elements of D_i are 1.34, 1.7, 1.4 and 1.

Thus,

$$D_i^{-1} N_i = \begin{bmatrix} .745 & .150 & .075 & .030 \\ 0 & .590 & .295 & .115 \\ 0 & 0 & .715 & .285 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

In equilibrium under constant external flows the fractions remaining in grade, promoted, and leaving the system are given by $\tilde{s}_i \hat{P}_i$, and from the above we have,

$$\tilde{s}_i \hat{P}_i = g_i^* D_i^{-1} N_i \hat{P}_i$$

$$= g_i^* \begin{bmatrix} .755 & .150 & .075 & .030 \\ 0 & .590 & .295 & .115 \\ 0 & 0 & .715 & .285 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .2 & .7 & .1 \\ .5 & .4 & .1 \\ .4 & .1 & .5 \\ 0 & .1 & .9 \end{bmatrix}$$

$$= g_i^* \begin{bmatrix} .255 & .590 & .155 \\ .403 & .288 & .309 \\ .286 & .100 & .714 \\ 0 & .100 & .900 \end{bmatrix}.$$

Because g_i^* can be any $1 \times w_i$ distribution vector, we see that the parameters of the combined process in equilibrium under constant external flows, $\tilde{s}_i \hat{P}_i$, must be a convex combination of the rows of the above matrix. The set of all such convex combinations has been previously designated H_i^*, and it is plotted with dashed lines in Figure 2. For long-range planning the planner should restrict his choice of parameters for the combined process to this set H_i^*.

Next let us assume that the planner estimates that in equilibrium the vector of the numbers entering the states in grade i would be approximately,

$$g_i = (30,\ 50,\ 20,\ 0).$$

Then the equilibrium stocks would be

$$\begin{aligned} s_i &= g_i\ N_i \\ &= (30,\ 60,\ 48,\ 19.2), \end{aligned}$$

$$\tilde{s}_i = (.191,\ .382,\ .305,\ .122),$$

and,

$$\tilde{s}_i\ \hat{P}_i = (.351,\ .329,\ .320).$$

The latter point is denoted by a triangle in Figure 2. In this case a convex combination of the points denoted by * and Δ in Figure 2 would seem a reasonable choice of the parameters of the combined process, i.e.,

$$(q_i,\ p_i,\ a_i) = \theta(.351,\ .329,\ .320) + (1.\theta)(.43,\ .43,\ .14),$$

$$\text{where } 0 \le \theta \le 1.$$

Small values of θ are used for short-range planning, and larger values of θ are used for long-range planning.

VII. AN APPLICATION TO RETRAINING PROBLEMS

A. INTRODUCTION

Consider an organization with people trained in various skills. Each person with a given skill belongs to a "skill group," which we will call "group" for simplicity. In this chapter we assume that the second characteristic in the state description is the group to which an individual belongs. We also assume that a person cannot belong to more than one group at a time.

Group membership may be quite explicit as in the case of the U.S. Marine Corps system of occupational fields or it can be implicit as in the case where group membership is determined by the number of years of formal education completed. In any case group membership defines a partition on the organization.

In many organizations retention and promotion vary considerably from group to group. Because of this it is often the case that people have to be retrained between groups in order to prevent surpluses and deficiencies of people in various skills. This is particularly true in the military enlisted personnel system. These retraining problems are the subject of this chapter.

We assume that there is a nonnegative cost associated with the retraining of a person from one group to another. The meaning and the numerical value of a retraining cost is left to the planner. For example, a planner may express his opinion that people in group k are unsuited for retraining

into group m by assigning a prohibitively high cost to such retraining.

Unlike the LOS or TIG models, in the retraining model we are quite interested in how many people change their second characteristic each period. We assume that assumptions A0, A1 and A2 of Chapter II still hold.

We are interested in long-range planning, so we begin with an equilibrium model, i.e., a model in which we assume that stocks and flows do not vary from one period to the next. The equilibrium model is very useful in determining achievable goals for an organization. Once determined, these goals can be used to judge short-range manpower policies. For example, use of the equilibrium model might show a long-range requirement for retraining into the sonarman group, and there might be a present surplus of sonarmen. The goals from the equilibrium model would cause us in this case to question a short-range recommendation to convert sonarmen training facilities to some other use.

In a retraining problem the planner would ideally like to specify the stocks and promotion rates for each grade/group combination (state) and then minimize retraining costs. Unfortunately, this is not generally possible, and the planner must compromise or trade off between desired stocks, desired promotion rates and minimal retraining costs. A major purpose of our development of the retraining model will be to show the close interaction between stocks, promotion rates and retraining costs. Consequently, we take a descriptive

rather than a prescriptive viewpoint of the retraining problem. The prescriptive approach would require the planner to specify a mathematical description of his preferences with respect to tradeoffs between stocks, promotion rates and retraining costs. Such a requirement is usually quite impracticable.

In the following development of the retraining model the attrition rates are treated as known and fixed. This may not be entirely acceptable. Retraining can be used as a method of reducing attrition. It can be used as an inducement to people to remain in the organization. This problem has not been treated here.

One purpose of the mathematical analysis of the retraining problem is to form the basis for an interactive computer program. Such a program would assist the planner in finding acceptable and feasible combinations of stocks, promotion rates and retraining costs. A device to be used in the interactive program, which is used in the mathematical analysis below, is proportionate control of stocks and promotion rates. For example, for a given grade the planner is required to specify the proportionate promotion rates for the various groups in that grade. This device may appear artificial; however it has great practical advantages. The U.S. Department of Defense is considering a promotion policy in which promotion rates within any grade must be equal for all groups. It is not our intention to join the debate over the efficacy of such a policy, but our results are quite useful in investigating the implications of such a policy.

B. DEFINITIONS AND THE BALANCE EQUATION

The organization is partitioned into states according to grade and group. It is assumed that transitions between states can be modeled by stationary fractional flows. The organization is assumed to be in equilibrium; stocks and flows do not change from one period to the next. One-period transitions from any grade are restricted to the same grade, the next higher grade or out of the system (see assumption A1).

There are K groups indexed by $k=1,\cdots,K$ and n grades indexed by $i=1,\cdots,n$. State (i,k) corresponds to grade i and group k. We define

s_{ik} = number of people in grade i and group k,

s_i = $(s_{i1},\cdots,s_{iK})$, a $1 \times K$ vector,

f_{ik} = number of people who enter the system in state (i,k) each period,

f_i = $(f_{i1},\cdots,f_{iK})$, a $1 \times K$ vector,

a_{ik} = fraction of those in state (i,k) at the end of one period who leave the system during the next period,

a_i = $[a_{i1},\cdots,a_{iK}]$, a $K \times 1$ vector,

q_{ik} = fraction of those in state (i,k) at the end of one period who are still in grade i (in any group) at the end of the next period,

p_{ik} = fraction of those in state (i,k) at the end of one period who are promoted to grade $i+1$ during the next period.

p_i = $[p_{i1},\cdots,p_{iK}]$, a $K \times 1$ vector.

From assumption A1 we have

$$q_{ik} + p_{ik} + a_{ik} = 1.$$

Nonnegative proportionality constants for the promotion rates, $\{\alpha_{ik}: i=1,\cdots,n,\ \ k=1,\cdots,K\}$, are specified so that in each grade i,

$$p_{ik} = p_i^* \alpha_{ik}, \quad k=1,\cdots,K, \tag{1}$$

where p_i^* is the maximum promotion rate over all groups in grade i.

Let,

$\alpha_i = [\alpha_{i1},\cdots,\alpha_{iK}]$, a K x 1 vector with maximum element equal to 1. Thus,

$$p_i = p_i^* \alpha_i , \tag{2}$$

and we see that whatever the promotion rates for the various groups in grade i may be, they will be proportioned according to the components of α_i.

We begin by considering all K groups in some grade i. The number of people entering grade i during any period is:

$$\text{No. into grade } i = \sum_{k=1}^{K} f_{ik} + \sum_{k=1}^{K} s_{i-1,k}p_{i-1,k}$$

$$= f_i\bar{1} + s_{i-1}p_{i-1} . \tag{3a}$$

The number of people leaving grade during any period is:

$$\text{No. out of grade } i = \sum_{k=1}^{K} s_{ik}a_{ik} + \sum_{k=1}^{K} s_{ik}p_{ik}$$

$$= s_i a_i + s_i p_i$$

$$= s_i(a_i + p_i). \tag{3b}$$

Under the assumption that the system is in equilibrium, the numbers entering and leaving a grade each period must be equal. Thus

$$f_i\bar{l} + s_{i-1}p_{i-1} = s_i(a_i + p_i). \tag{4}$$

The above equation will be referred to as the balance equation for grade i.

C. PARAMETRIC CALCULATIONS FOR PROMOTION RATES

For specified stocks we use the proportionality constants for the promotion rates to solve for the promotion rates one grade at a time starting with grade 1.

$$\begin{aligned} f_1\bar{l} &= s_1(a_1 = p_1) \\ &= s_1a_1 + s_1(p_1^*\alpha_1) \\ &= s_ia_1 + p_1^*s_1\alpha_1 \end{aligned}$$

(Recall the p_i^*'s are scalars.)

Thus,

$$p_1^* = \frac{f_1\bar{l} - s_1a_1}{s_1\alpha_1},$$

and

$$p_1 = p_1^*\alpha_1 .$$

For grades i=2,···,n:

$$\begin{aligned} f_i\bar{l} + s_{i-1}p_{i-1} &= s_i(a_i + p_i) \\ &= s_ia_i + p_i^*s_i\alpha_i . \end{aligned}$$

Thus,

$$p_i^* = \frac{f_i\bar{l} + s_{i-1}p_{i-1} - s_ia_i}{s_i\alpha_i} \tag{5}$$

and,

$$p_i = p_i^* \alpha_i \, .$$

Note that when the stocks are specified we can solve for all the promotion rates without explicitly considering the retraining flows between groups.

We note that:

1) $f_i \bar{1}$ is the number of people who enter grade i each period from outside the system,

2) $s_{i-1} p_{i-1}$ is the number of people promoted into grade i (from grade i-1) each period,

3) $s_i a_i$ is the number of people who leave the system from grade i each period. Thus $p_i^* \geq 0$ if and only if the total number entering grade i each period is no smaller than the number leaving the system from grade i each period.

Feasible promotion rates must satisfy $0 \leq p_{ik} \leq 1 - a_{ik}$. The values of p_{ik} computed above from the balance equations may not satisfy these constraints. It may also be the case that the promotion rates computed above are unacceptable for practical reasons. We next consider how one may trade off between stocks and promotion rates to obtain a satisfactory feasible set of rates.

D. PARAMETRIC CALCULATIONS FOR STOCKS

Let us suppose that the specified stocks yield infeasible promotion rates in grade i, e.g., $p_i^* < 0$. We might then approach the problem from the opposite direction in grade i by specifying the promotion rates and solving the balance equations for the stocks. The difficulty is that the solution

to such a problem is not unique. We will discuss two parametric methods that lead to unique solutions for the stocks in grade i when the values of the promotion rates in grade i are specified.

In the context of interactive programming the parametric calculation of stocks is used to make adjustments when the initially specified stocks have led to infeasible or unacceptable promotion rates. In Method 1 a specified proportionality between stocks in the various groups is maintained. In Method 2 we maintain a specified proportionality between the deviations of the stocks from their desired levels.

1. Method 1

In this method we specify nonnegative proportionality constants for the stocks in the various groups in grade i, $\{\gamma_{ik}:k=1,\cdots,K\}$ where $\sum_{k=1}^{K}\gamma_{ik} = 1$. Then,

$$s_{ik} = s_i^*\gamma_{ik} \,. \tag{6}$$

where s_i^* is the total number of people in grade i.

Let,

$$\gamma_i = (\gamma_{i1},\cdots,\gamma_{iK}), \text{ a } 1 \times K \text{ vector.}$$

Then,

$$s_i = s_i^*\gamma_i \,. \tag{7}$$

Let the promotion rate vector for grade i, p_i, be specified. The balance equation (4) for grade i is

$$f_i\bar{1} + s_{i-1}p_{i-1} = s_i(a_i + p_i)$$

$$= s_i^*\gamma_i(a_i + p_i) \,.$$

Thus,

$$s_i^* = \frac{f_i\bar{l} + s_{i-1}p_{i-1}}{\gamma_i(a_i + p_i)} \tag{8}$$

and,

$$s_{ik} = s_i^*\gamma_{ik}.$$

To interpret these results note that,

$$\gamma_i(a_i + p_i) = \frac{s_i^*\gamma_i(a_i + p_i)}{s_i^*}$$

$$= \frac{s_i(a_i + p_i)}{s_i\bar{l}}$$

= fraction of those in grade i who leave grade i (by attrition or promotion) each period.

2. Method 2

In this method we specify proportionality constants for the changes in the stocks in grade i, $\{\delta_{ik};\ k=1,\cdots,K\}$. We suppose that the desired stocks in grade i, denoted by $\{\hat{s}_{ik};\ h=1,\cdots,K\}$ lead to infeasible or unacceptable promotion rates, and require that the stocks satisfy not only the balance equation but also the parametric constraints,

$$s_{ik} = (1 + c_i\ \delta_{ik})\hat{s}_{ik}$$

where c_i is a scalar to be determined.

Let,

$\delta_i = (\delta_{i1},\cdots,\ \delta_{iK})$, a 1 x K vector,

$\hat{s}_i = (\hat{s}_{i1},\cdots,\hat{s}_{iK})$, a 1 x K vector,

and define an operator @ such that $E^m \times E^m \rightarrow E^m$ and

$\delta_i@\hat{s}_i = (\delta_{i1}\hat{s}_{i1},\cdots,\ \delta_{iK}\hat{s}_{iK})$. (See Appendix A.)

We then have

$$s_i = \hat{s}_i + c_i \ \delta_i @ \hat{s}_i \, .$$

Using Method 2 parameterization the balance equation (4) for grade i is

$$\begin{aligned} f_i \bar{l} + s_{i-1} p_{i-1} &= s_i (a_i + p_i) \\ &= (\hat{s}_i + c_i \ \delta_i @ \hat{s}_i)(a_i + p_i) \, . \end{aligned}$$

Thus,

$$c_i = \frac{f_i \bar{l} + s_{i-1} p_{i-1} - \hat{s}_i (a_i + p_i)}{(\delta_i @ \hat{s}_i)(a_i + p_i)}$$

and,

$$s_{ik} = (1 + c_i \ \delta_{ik}) \hat{s}_{ik} \, .$$

One advantage of Method 2 is that the stock in some group k can be held constant at the desired level $\hat{s}_{ik}$ by assigning δ_{ik} a value of zero.

A significant difference between the two method of parameterizing the stocks is that Method 1 always yields nonnegative stocks while Method 2 may not. In the initial stages of an investigation of retraining policies, Method 1 would seem more practical. After feasible and not-too-unacceptable policies have been derived, then Method 2 could be used to advantage. Only Method 1 will be used to parameterize stocks in subsequent sections.

E. RETRAINING FLOWS

Once we have arrived at a set of stocks and promotion rates that are acceptable, we may compute the required retraining flows. Let,

r_{ik} = number of people retrained out of (>0) or into (<0) state (i,k) each period

r_i = $(r_{i1},\cdots,r_{iK})$, a 1 x K vector.

We derive an expression for the retraining flow r_{ik} by considering the numbers of people that enter and leave state (i,k) each period in the absence of retraining:

No. into state (i,k) = $f_{ik} + s_{i-1,k}p_{i-1,k}$

No. out of state (i,k) = $s_{ik}(a_{ik} + p_{ik})$.

Under the assumption that the system is in equilibrium, the flows into and out of state (i,k) must be equal; we use the retraining flow r_{ik} to bring about this equality:

$$r_{ik} = (f_{ik} + s_{i-1,k}p_{i-1,k}) - s_{ik}(a_{ik} + p_{ik}) . \qquad (9)$$

The balance equation (4) ensures that

$$\sum_{k=1}^{K} r_{ik} = 0,$$

i.e., everyone retrained out of one group is retrained into another group.

F. THE RETRAINING TRANSPORTATION PROBLEM

We assume there is a known cost $c_i(k,m)$ for retraining a person in grade i from group k to group m. We assume that all costs, including $c_i(k,k)$, are strictly positive.

For any set of stocks and promotion rates in grade i that satisfy the balance equation, we may compute the retraining flows, $\{r_{ik}\}$, and treat these as supplies of people for retraining ($r_{ik} \geq 0$) and demands for retrained people ($r_{ik} < 0$). The balance equation ensures that total supply equals total demand.

Let,

$x_i(k,m)$ = number of people retrained from state (i,k) to state (i,m) each period.

In order to match supplies and demands at minimal cost, we must solve the following linear program.

$$\text{P1]}\quad \min \sum_{h=1}^{K} \sum_{m=1}^{K} c_i(k,m)\ x_i(k,m)$$

$$\text{ST}\quad \sum_{m=1}^{K} (x_i(k,m) - x_i(m,k)) = r_{ik};\ k=1,\cdots,K$$

$$x_i(k,m) \geq 0;\ k=1,\cdots,K \quad m=1,\cdots,K.$$

The equality constraints in P1 ensure that for each state (i,k) the net number of people retrained out of or into state (i,k) matches the supply of or demand for retrained people for that state.

We note that,

$$\sum_{m=1}^{K} (x_i(k,m) - x_i(m,k)) = \sum_{\substack{m=1 \\ m\neq k}}^{K} (x_i(k,m) - x_i(m,k))$$

is an identity. That is, $x_i(k,k)$ does not appear in the equality constraints of P1. The nonnegativity constraints do require $x_i(k,k) \geq 0$. Because we assume $c_i(k,k) > 0$, the optimal solution to P1 will always have $x_i(k,k) = 0$.

The dual of P1 is denoted D1:

$$\text{D1]}\quad \max \sum_{k=1}^{K} r_{ik}\ v_{ik}$$

$$\text{ST}\quad v_{ik} - v_{im} \leq c_i(k,m);\ k=1,\cdots,K \quad m=1,\cdots,K.$$

$$v_{ik}\ \text{UNRESTRICTED},\ k=1,\cdots,K.$$

It will be convenient to define the following vectors:

$$x_i = [x_i(1,1),\cdots,x_i(1,K),\ x_i(2,1),\cdots,x_i(K,1),\cdots,x_i(K,K)],$$

a $K^2 \times 1$ vector,

$v_i = [v_{i1},\cdots,v_{ik}]$, a $K \times 1$ vector.

The linear program P1 is quite similar to the classical transportation problem. It becomes a transportation problem if we make the following assumption.

Assumption A3. For any $h,j,m \varepsilon \{1,\cdots,K\}$,

$$c_i(k,m) < c_i(k,j) + c_i(j,m).$$

The practical implication of this assumption is that it must be cheaper to retrain a person from group k to group m than it is to retrain one person from group k to group j and retrain another person from group j to group m.

The theoretical implication of this assumption is that the optimal values of some readily identified variables in P1 must be zero.

Let,

$$K_i^+ = \{k\colon r_{ik} \geq 0\},$$

$$K_i^- = \{k\colon r_{ik} < 0\}.$$

Consider the case in which $j \varepsilon K_i^+$, the optimal solution to D1 is v_i^0, and the optimal solution to P1 is x_i^0.

Suppose that $x_i^0(k,j) > 0$. By complementary slackness it must then be true that,

$$v_{ik}^0 - v_{ij}^0 = c_i(k,j). \tag{10}$$

But $j \varepsilon K_i^+$, so there must be some m such that $x_i^o(j,m) > 0$. This implies by complementary slackness that,

$$v_{ij}^o - v_{im}^o = c_i(j,m). \qquad (11)$$

Adding (10) and (11) and using assumption A3, we have

$$v_{ik}^o - v_{im}^o = c_i(k,j) + c_i(j,m) > c_i(k,m),$$

implying that the optimal solution to D1, v_i^o, is not feasible, a contradiction. Consequently, it cannot be true that $x_i^o(k,j) > 0$ when $j \varepsilon K_i^+$. By a similar argument, we can show that it cannot be true that $x_i^o(j,m) > 0$ when $j \varepsilon K_i^-$.

Thus assumption A3 implies that in the optimal primal solution, if $x_i^o(k,m) > 0$, then $k \varepsilon K_i^+$, $m \varepsilon K_i^-$. Removing the variables which must be zero in any optimal solution from P1 leads to the primal and dual equivalents of P1 and D1 under assumption A3:

PRTP] min $\displaystyle\sum_{k \varepsilon K_i^+} \sum_{m \varepsilon K_i} c_i(k,m)\ x_i(k,m)$

ST $\displaystyle\sum_{m \varepsilon K_i^-} x_i(k,m) = r_{ik},\ k \varepsilon K_i^+$

$\displaystyle\sum_{k \varepsilon K_i^+} - x_i(k,m) = r_{im},\ m \varepsilon K_i^-$

$x_i(k,m) \geq 0.$

DRTP] max $\displaystyle\sum_{k=1}^{K} r_{ik}\ v_{ik}$

ST $v_{ik} - v_{im} \leq c_i(k,m);\ k \varepsilon K_i^+,\ m \varepsilon K_i^-$

v_{ik} UNRESTRICTED.

The abbreviations PRTP and DRTP denote the primal and dual retraining transportation problems respectively. The programs PRTP and DRTP are in the form of the classical transportation problem.

We should pause to note some of the assumptions inherent in modeling retraining costs in this rather simple manner.

1) There is no provision for set up costs.

2) Each grade is treated as a separate problem.

3) Marginal retraining costs are assumed constant with respect to the number retrained.

The computational tractability of the simple transportation model is no small consideration, and one suspects that the model is adequate for the purposes at hand, viz., long-range planning.

Let the optimal value and the optimal solution to DRTP be denoted d_i^o and $v_i^o = [v_{i1}^o, \cdots, v_{iK}^o]$ respectively. Then,

$$d_i^o = \sum_{k=1}^{K} r_{ik} \, v_{ik}^o \, . \tag{12}$$

We note that the optimal solution to DRTP is unique only to an additive constant.

G. RELATION OF RETRAINING COSTS TO STOCKS AND PROMOTION RATES

In this section we assume that the stocks and promotion rates in grade i-1 have been determined, and we consider how our choice of the stocks and promotion rates in grade i will affect the optimal retraining costs in that grade.

We will parameterize both the stocks and the promotion rates in grade i using Method 1 for the parametric calculation

of the stocks. Thus, we must specify two sets of proportionality constants (see equations (2) and (7)):

1) $\{\alpha_{ik}, k=1,\cdots,K\}$ for the promotion rates,

2) $\{\gamma_{ik}, k=1,\cdots,K\}$ for the stocks.

We consider the case in which the total number of people entering grade i each period is known and fixed:

$$f_i + s_{i-1} @ p_{i-1} = g_i, \text{ a known } 1 \times K \text{ vector}$$

$$f_i\bar{1} + s_{i-1}p_{i-1} = g_i\bar{1}, \text{ a scalar.}$$

The stocks and promotion rates for the various groups in grade i are determined by the choice of s_i^* and p_i^* :

$$s_{ik} = s_i^* \gamma_{ik} ,$$

$$p_{ik} = p_i^* \alpha_{ik} .$$

The balance equation (4) places a constraint on the choice of s_i^* and p_i^*. We consider s_i^* as the independent variable, so the balance equation,

$$\begin{aligned} g_i\bar{1} &= s_i(a_i + p_i) \\ &= s_i^*\gamma_i(a_i + p_i^*\alpha_i), \end{aligned}$$

leads to the relation of the dependent variable p_i^* to the independent variable s_i^* :

$$p_i^* = \frac{g_i\bar{1} - s_i^* \gamma_i a_i}{s_i^* \gamma_i \alpha_i} . \tag{13}$$

This relation is sketched on the following page.

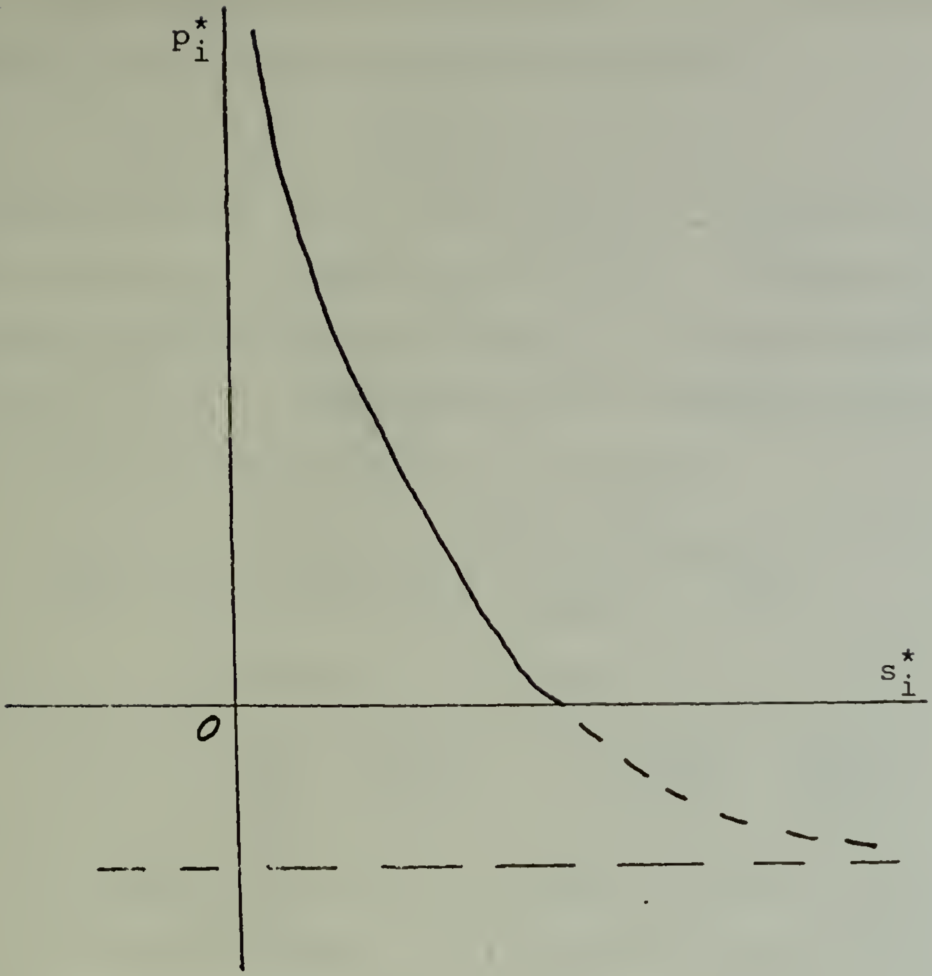

Figure 2. Sketch of s_i^* and p_i^*

The curve is asymptotic to the lines $s_i^* = 0$ and $p_i^* = -\gamma_i a_i/\gamma_i \alpha_i$.

We must have $p_i^* \geq 0$, which by (13) is equivalent to $s_i^* \gamma_i a_i \leq g_i \bar{l}$, i.e., the total number leaving the system from grade i each period must be no greater than the total number entering grade i each period. It is only when the above inequality is strict that we may (indeed must) promote people out of grade i.

Let us suppose that the couple (s_i^*, p_i^*) satisfies (13). We then use the retraining flow equation,

$$r_{ik} = g_{ik} - s_i^* \gamma_{ik}(a_{ik} + p_i^* \alpha_{ik}), \qquad (9)$$

to determine the RHS of the retraining transportation problem. The value of r_{ik} depends on s_i^* and p_i^*. Because p_i^* is uniquely determined by s_i^* from (13), we may consider r_{ik} a function of s_i^*. Substituting the expression for p_i^* from (13) into the equation for r_{ik} we have,

$$r_{ik}(s_i^*) = g_{ik} - s_i^*\gamma_{ik}(a_{ik} + (\frac{g_i\bar{l} - s_i^*\gamma_i a_i}{s_i^*\gamma_i\alpha_i})\alpha_{ik}).$$

After some manipulation this leads to,

$$r_{ik}(s_i^*) = g_i\bar{l}(\frac{g_{ik}}{g_i\bar{l}} - \frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i}) + s_i^*\gamma_i a_i(\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} - \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i}). \qquad (14)$$

We assume the vector of group-to-group retraining costs, c_i, is fixed and known, so the optimal value of the retraining transportation problem is a function of the vector of retraining flows, r_i. But we have shown that for fixed flow into grade i, g_i, the retraining flows are a function of s_i^*, so the optimal value of the retraining transportation problem is a function of s_i^*. We show this by rewriting the primal and dual of the retraining transportation problem.

$$\text{PRTP]} \quad \min \phi_i = \sum_{k\varepsilon K_i^+} \sum_{m\varepsilon K_i^-} c_i(k,m)x_i(k,m)$$

$$\text{ST} \quad \sum_{m\varepsilon K_i^-} x_i(k,m) = r_{ik}(s_i^*), \quad k\varepsilon K_i^+$$

$$\sum_{k\varepsilon K_i^+} x_i(k,m) = -r_{im}(s_i^*), \quad m\varepsilon K_i^-$$

$$x_i(k,m) \geq 0.$$

$$\text{DRTP]} \quad \max d_i(s_i^*) = \sum_{k=1}^{K} r_{ik}(s_i^*)v_{ik}$$

$$\text{ST} \quad v_{ik} - v_{im} \leq C_i\ (k,m);\ k\varepsilon K_i^+,\ m\varepsilon K_i^-$$

$$v_{ik},\ v_{im}\ \text{UNRESTRICTED}.$$

The question that we address in this section is how the optimal retraining costs vary with respect to the total stocks s_i^*.

For some specified values of s_i^* let $d_i^o(s_i^*)$ be the optimal value of the retraining costs and let $v_i^o(s_i^*)$ be the optimal solution to the dual retraining transportation problem.

We then have,

$$\frac{\partial d_i^o(s_i^*)}{\partial s_i^*} = \frac{\partial}{s_i^*} \sum_{k=1}^{K} r_{ik}(s_i^*)v_{ik}^o(s_i^*)$$

$$= \sum_{k=1}^{K} (v_{ik}^o(s_i^*) \frac{\partial r_{ik}(s_i^*)}{\partial s_i^*} + r_{ik}(s_i^*) \frac{\partial v_{ik}^o(s_i^*)}{\partial s_i^*}.$$

From (14) we have immediately,

$$\frac{\partial r_{ik}(s_i^*)}{\partial s_i^*} = \gamma_i a_i\ (\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} - \frac{\gamma_{ik}a_{ik}}{\gamma_i\alpha_i}).$$

The question then is, how does the dual optimal solution $v_i^o(s_i^*)$ vary with s_i^*. We note in the dual that varying s_i^* does not alter the feasibility of $v_i^o(s_i^*)$. We note in the primal that varying s_i^* varies the RHS of the program, so the primal optimal solution must vary with s_i^*. But suppose that the primal basis that was optimal at s_i^* is still feasible at $s_i^* + \varepsilon$ for $|\varepsilon|$ sufficiently small. Then this primal basis

is feasible, it satisfies the complementary slackness condition with respect to $v_i^o(s_i^*)$, and $v_i^o(s_i^*)$ is feasible in the dual, so $v_i^o(s_i^*)$ is still the dual optimal solution. In summary, the dual optimal solution, $v_i^o(s_i^*)$, does not change with s_i^* so long as the primal basis which was optimal at s_i^* remains feasible. (See Dantzig, 1963.)

Consequently, if for some $\varepsilon>0$ changing the stocks from s_i^* to $s_i^* \pm \varepsilon$ does not cause a primal basis change, then,

$$\frac{\partial v_{ik}^o(s_i^*)}{\partial s_i^*} = 0 \,,$$

and

$$\frac{\partial d_i^o(s_i^*)}{\partial s_i^*} = \gamma_i a_i \sum_{k=1}^{K} v_{ik}^o(s_i^*) \left(\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} - \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i}\right) \tag{15}$$

We note that:

1) $\gamma_i a_i = \dfrac{s_i^*\gamma_i a_i}{s_i^*} = \dfrac{s_i a_i}{s_i\bar{1}}$

= fraction of those in grade i who leave the system each period

2) $\dfrac{\gamma_{ik}a_{ik}}{\gamma_i a_i} = \dfrac{s_i^*\gamma_{ik}a_{ik}}{s_i^*\gamma_i a_i} = \dfrac{s_{ik}a_{ik}}{s_i a_i}$

= fraction of those leaving system from grade i who leave from state (i,k)

3) $\dfrac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} = \dfrac{s_i^*\gamma_{ik}p_i^*\alpha_{ik}}{s_i^*\gamma_i p_i^*\alpha_i} = \dfrac{s_{ik}p_{ik}}{s_i p_i}$

= fraction of those promoted from grade i who are promoted from state (i,k).

1. Special Case 1 -- Costs Constant

Special case 1 is that in which for some positive constant c,

$$\alpha_{ik} = ca_{ik}, \quad k=1,\cdots,K.$$

In this special case,

$$\alpha_i = ca_i$$

$$\frac{\alpha_{ik}}{\gamma_i \alpha_i} = \frac{ca_{ik}}{c\gamma_i a_i} = \frac{a_{ik}}{\gamma_i a_i},$$

$$\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i \alpha_i} - \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i} = 0,$$

$$r_{ik}(s_i^*) = g_i\bar{1}\left(\frac{g_{ik}}{g_i\bar{1}} - \frac{\gamma_{ik}\alpha_{ik}}{\gamma_i \alpha_i}\right),$$

$$\frac{\partial d_i^*(s_i^*)}{\partial s_i^*} = 0 .$$

Thus, when promotion rates are proportional to attrition rates, the retraining flows do not depend on the total stocks s_i^*. Consequently, the optimal retraining costs do not depend on s_i^*.

The sensitivity of retraining costs to the total stocks depends in some sense on how promotion and attrition rates depart from the foregoing special case.

2. Special Case 2 -- Costs Increase with Stocks

A second special case is that in which for each $k\varepsilon\{1,\cdots,K\}$, either,

1) $\dfrac{g_{ik}}{g_i\bar{1}} \le \dfrac{\gamma_{ik}\alpha_{ik}}{\gamma_i \alpha_i} \le \dfrac{\gamma_{ik}a_{ik}}{\gamma_i a_i}$,

or,

$$2)\quad \frac{g_{ik}}{g_i\bar{1}} \geq \frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} \geq \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i} .$$

We note that,

$\frac{g_{ik}}{g_i\bar{1}}$ = fraction of those entering grade i who enter in group k.

Thus, condition 1) above indicates that the fraction of grade i "entrants" entering group k is smaller than the fraction of grade i "promotees" promoted from group k which in turn is smaller than the fraction of grade i "leavers" leaving from group k. From equation (14), condition 1) implies that both $r_{ik}(s_i^*)$ and its partial derivative with respect to s_i^* are nonpositive. That is, there is a demand for retrained people in group k of grade i, and as the total stocks in grade i increase so does this demand for retrained people in group k.

Condition 2) is simply condition 1) with the inequalities reversed. Condition 2) implies that both $r_{ik}(s_i^*)$ and its partial derivative with respect to s_i^* are nonnegative. So there is an excess of people in group k of grade i, and as the total stocks in grade i increase so does this excess of people in group k.

In this special case increasing total stocks increases the supplies in the groups that have people available for retraining and increases the requirements in the groups that have a need for retrained people. Consequently, increasing total stocks, s_i^*, increases the optimal retraining costs.

3. Special Case 3 -- Costs Decrease with Stocks

A third special case is that in which for some values of s_i^* we have for each $k\varepsilon\{1,\cdots,K\}$, either,

$$1)\quad r_{ik}(s_i^*) > 0 \text{ and } \frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} \leq \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i},$$

or,

$$2)\quad r_{ik}(s_i^*) < 0 \text{ and } \frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} \geq \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i}.$$

Condition 1) above indicates that group k has people available for retraining, and the fraction of grade i "promotees" promoted from group k is smaller than the fraction of grade i "leavers" leaving from group k. If the total stocks, s_i^*, are increased then the number leaving the system from group k increases in direct proportion to s_i^*. However, increasing s_i^* causes the overall promotion rate p_i^* to decrease, so the number promoted from group k does not increase as rapidly as s_i^*. Under condition 1) we then have a decrease in the number to be retrained when s_i^* increases.

Condition 2) is simply condition 1) with the inequalities reversed. Under condition 2) there is a requirement for retrained people, but this requirement decreases as s_i^* increases.

In this special case increasing total stocks decreases the supplies in the groups that have people available for retraining and decreases the requirements in the groups that have a need for retrained people. Consequently, increasing total stocks, s_i^*, decreases the optimal retraining costs.

The rate of change of retraining costs with respect to promotion rates may be computed from

$$\frac{\partial s_i^*}{\partial p_i^*} = - \frac{s_i^* \gamma_i \alpha_i}{\gamma_i (a_i + p_i^* \alpha_1)}$$

and,

$$\frac{\partial d_i^o}{\partial p_i^*} = \frac{\partial d_i^o}{\partial s_i^*} \cdot \frac{\partial s_i^*}{\partial p_i^*} .$$

H. TWO-GRADE OPTIMIZATION OF RETRAINING COSTS

In this section we develop a technique for jointly varying the stocks and promotion rates in two grades, i and i+1, in such a way that the stock and promotion rates in grades below grade i and above grade i+1 need not be changed. We then show how the retraining costs in grades i and i+1 respond to these stock and promotion rate changes.

Recall the balance equations (4) for grades i, i+1, and i+2:

i) $f_i \bar{l} + s_{i-1} p_{i-1} = s_i^* \gamma_i (a_i + p_i^* \alpha_i) = s_i^* (\gamma_i a_i) + s_i^* p_i^* (\gamma_i \alpha_i)$

i+1) $f_{i+1} \bar{l} + s_i^* p_i^* \gamma_i \alpha_i = s_{i+1}^* \gamma_{i+1} (a_{i+1} + p_{i+1}^* \alpha_{i+1})$

i+2) $f_{i+2} \bar{l} + s_{i+1}^* p_{i+1}^* \gamma_{i+1} \alpha_{i+1} = s_{i+2} (a_{i+2} + p_{i+2})$:

We consider s_{i-1} and p_{i-1} as fixed vectors. If s_i^* is varied, then p_i^* must also change in order to satisfy i). Furthermore, the product $s_i^* p_i^*$ cannot remain constant as s_i^* is varied. This means that satisfying i+1) when s_i^* is varied will require a change in s_{i+1}^* or p_{i+1}^* or both. Suppose that as s_i^* is

varied we change the value of s^*_{i+1} so that i+1) is satisfied and the produce $s^*_{i+1}p^*_{i+1}$ is held constant. This is possible and the values of s^*_{i+1} and p^*_{i+1} are uniquely determined. By holding $s^*_{i+1}p^*_{i+1}$ constant equation i+2) is satisfied without changing stocks in grade i+2. Thus we have restricted the effects of varying s^*_i to grades i and i+1; all other grades are unaffected.

A matter of interest is how the retraining costs vary as a consequence of changing stocks by the foregoing technique. We have shown in the previous section how retraining costs in grade i respond to changes in stocks when the input flow is fixed. We next consider how retraining costs in grade i+1 respond to changes in s^*_i when the promotion flow from grade i+1, $s^*_{i+1}p^*_{i+1}\gamma_{i+1}\alpha_{i+1}$, is held constant.

We use three equations:

1) Balance equation

$$f_{i+1}\bar{1} + s^*_i p^*_i \gamma_i \alpha_i = s^*_{i+1}\gamma_{i+1}(a_{i+1}+p^*_{i+1}\alpha_{i+1}) \qquad (4')$$

2) Retraining flow equation

$$r_{i+1,k} = f_{i+1,k} + s^*_i p^*_i \gamma_{ik}\alpha_{ik} - s^*_{i+1}\gamma_{i+1,k}\,(a_{i+1,k}+p^*_{i+1}\alpha_{i+1,k}) \qquad (9')$$

3) Fixed promotion flow constraint

$$s^*_{i+1}p^*_{i+1} = \text{constant}.$$

Taking the partial derivative of the balance equation with respect to s^*_i yields,

$$\frac{\partial s^*_{i+1}}{\partial s^*_i} = -\frac{\gamma_i a_i}{\gamma_{i+1}a_{i+1}}.$$

The partial derivative of total stocks in grade i+1 with respect to total stocks in grade i is negative when $s_{i-1}^{*}p_{i-1}^{*}$ and $s_{i+1}^{*}p_{i+1}^{*}$ are held constant. This may be explained in terms of the promotion rate in grade i. Decreasing p_i^{*} causes an increase in s_i^{*} but a decrease in $s_i^{*}p_i^{*}$. This in turn causes a decrease in s_{i+1}^{*}. Taking the partial derivative of the retraining flow equation then leads to,

$$\frac{\partial r_{i+1,k}}{\partial s_i^{*}} = \gamma_i a_i \left(\frac{\gamma_{i+1,k}\, a_{i+1,k}}{\gamma_{i+1}\, a_{i+1}} - \frac{\gamma_{ik}\, \alpha_{ik}}{\gamma_i \alpha_i}\right).$$

We then have,

$$\frac{\partial d_{i+1}^{o}}{\partial s_i^{*}} = \gamma_i a_i \sum_{k=1}^{K} v_{i+1,k}^{o} \left(\frac{\gamma_{i+1,k}\, a_{i+1,k}}{\gamma_{i+1}\, a_{i+1}} - \frac{\gamma_{ik}\, \alpha_{ik}}{\gamma_i \alpha_i}\right)$$

and,

$$\frac{\partial (d_i^{o} + d_{i+1}^{o})}{\partial s_i^{*}} = \gamma_i a_i \sum_{k=1}^{K} \Big[-v_{ik}^{o} \frac{\gamma_{ik}\, a_{ik}}{\gamma_i a_i} + (v_{ik}^{o} - v_{i+1,k}^{o}) \frac{\gamma_{ik} \alpha_{ik}}{\gamma_i \alpha_i} + v_{i+1,k}^{o} \frac{\gamma_{i+1,k}\, a_{i+1,k}}{\gamma_{i+1}\, a_{i+1}} \Big]. \qquad (16)$$

This equation provides a rather practical means of guiding the planner in his search for an acceptable combination of stocks, promotion rates and retraining costs. The scalar coefficients of the optimal dual variables in (16) are readily computed, and they can be stored in a 3 x K matrix. For the value of s_i^{*} specified by the planner the optimal dual variables for the corresponding retraining transportation problem are found and (16) is computed using 3K multiplications and additions. The sign of the partial derivative in (16) indicates to the planner how he might reduce retraining

costs. Whether the planner chooses to make the indicated change in total stocks to reduce costs will depend on how he perceives his constraints on stocks and promotion rates. At any rate he may change the stocks in grades i and i+1 so as to decrease retraining costs and confine the effects of the stock changes to grades i and i+1.

A special case of some practical interest is that in which retraining costs,

1) are the same in grades i and i+1, i.e.,

$$c_i(k,m) = c_{i+1}(k,m),$$

and,

2) depend only on the group retrained into, i.e.,

$$c_i(k,m) = c_i(m).$$

In this special case any primal feasible solution to the retraining transportation problem is optimal, and an optimal solution to the dual is,

$$\begin{aligned} v^o_{ik} &= 0 && \text{if } k \varepsilon K_i^+ \\ &= -c_i(k) && \text{if } k \varepsilon K_i^- . \end{aligned}$$

One then has,

$$\frac{\partial(d_i^o + d_{i+1}^o)}{\partial s_i^*} = \gamma_i a_i \sum_{k \varepsilon K_i^-} c_i(k) \left(\frac{\gamma_{ik}\, a_{i+1}}{\gamma_i a_i} - \frac{\gamma_{i+1,k}\, a_{i+1,k}}{\gamma_{i+1}\, a_{i+1}}\right).$$

Roughly speaking, the above equation compares attrition rates in grades i and i+1 weighted by retraining costs, and indicates that costs will be reduced by reducing stocks in the grade that has the greater weighted attrition rate.

I. THE RELATION OF TRAINING COSTS TO TOTAL FLOWS

In Sections G and H we studied the behavior of the optimal retraining costs when the total stock in grade i, s_i^*, is the

independent variable. This approach has definite practical appeal because s_i^* is likely to be of importance to the planner. In this section we shall consider the behavior of the optimal retraining costs when there are two independent variables: the total number of people entering grade i each period and the total number of people promoted out of grade i each period. The two scalar variables will be referred to as the "total flow into" and the "total flow out of" grade i. It should be noted, that the "total flow out of" grade i does not include the number of people that leave the system from grade i each period.

We begin with an assumption about the external flows.

Assumption. For grades $i=2,\cdots,n$ the external flows are zero vectors, i.e., $f_i = \bar{0}$, $i=2,\cdots,n$. There is a known distribution vector $\tilde{f}_1$ such that the external flow for grade 1 is always a scalar multiple of $\tilde{f}_1$, i.e., f_1 always satisfies $\frac{f_1}{f_1\bar{1}} = \tilde{f}_1$.

The practical implications of this assumption are that hiring is restricted to grade 1, and the vector of people hired in grade 1 has a known distribution. The theoretical implications of this assumption are that the total flow out of grade i is the total flow into grade i+1, and the distribution of any total flow is specified by $\tilde{f}_1$ or $(\gamma_i @ \alpha_i)/\gamma_i\alpha_i$. (The latter point will be explained.)

Let us define a 1 x K vector $g_i=(g_{i1},\cdots,g_{iK})$ by

$$g_i = f_1 \quad \text{when } i=1$$

$$= s_{i-1}@p_{i-1}, \quad \text{when } i=2,\cdots,n.$$

We may also write g_i as the product of a total flow and a distribution vector,

$$g_i = (f_1\bar{1})\ \tilde{f}_1, \quad \text{when } i=1$$

$$= (s^*_{i-1}\ p^*_{i-1}\gamma_{i-1}\alpha_{i-1})\ \frac{\gamma_{i-1}@\alpha_{i-1}}{\gamma_{i-1}\ \alpha_{i-1}}, \quad \text{when } i=2,\cdots,n\ .$$

where,

$f_1\bar{1}$ = total flow into grade 1, (a scalar)

$s^*_{i-1}p^*_{i-1}\ \gamma_{i-1}\alpha_{i-1}$ = total flow into grade i, $i=2,\cdots,n$,

(a scalar), and the vectors $\tilde{f}_1$ and $\gamma_{i-1}@\alpha_{i-1}/\gamma_{i-1}\alpha_{i-1}$ are distribution vectors. Thus the vector of the numbers of people entering the various groups in grade i, g_i, is always equal to a scalar times a known distribution vector.

By our definitions of "total flow into" and "total flow out of" any grade we have,

$g_i\bar{1}$ = total flow into grade i, $i=1,\cdots,n$

= total flow out of grade i-1, $i=2,\cdots,n$.

Note that the total flow out of grade n must be zero, so we define $g_{n+1} = \bar{0}$.

Notation is simplified somewhat by defining,

$I_i = g_i\bar{1}$, a scalar,

$0_i = g_{i+1}\bar{1}$, a scalar.

The mnenomic I for "in" and 0 for "out" is intended.

We define the "flow plane" for grade i as the non-negative quadrant of E^2 with points having coordinates $(I_i, 0_i)$. That is, the total flow into grade i, I_i, and the total flow out of grade i, 0_i, define a point in the flow plane for grade i.

Lemma 1. For any point $(I_i, 0_i)$ in the flow plane for grade i the values of s_i^* and p_i^* that satisfy the balance equation are unique. In particular,

$$s_i^* = \frac{I_i - 0_i}{\gamma_i a_i} , \tag{17}$$

$$p_i^* = \left(\frac{\gamma_i a_i}{\gamma_i \alpha_i}\right)\left(\frac{0_i}{I_i - 0_i}\right) = \frac{0_i}{s_i^* \gamma_i \alpha_i} . \tag{18}$$

Proof. Recall that the balance equation (4) is,

$$f_i \bar{I} + s_{i-1} p_{i-1} = s_i (a_i + p_i) .$$

But under the assumption on external flows,

$$f_i \bar{I} + s_{i-1} p_{i-1} = I_i ,$$

so the balance equation can be written,

$$I_i = s_i a_i + s_i p_i .$$

But,

$$\begin{aligned} s_i p_i &= s_i^* p_i^* \gamma_i \alpha_i \\ &= 0_i , \end{aligned}$$

and,

$$s_i a_i = s_i^* \gamma_i a_i .$$

Thus,

$$I_i = s_i^* \gamma_i a_i + 0_i .$$

Rearranging terms we have,

$$s_i^* = (I_i - 0_i)/\gamma_i a_i \, .$$

The total flow out of grade i is

$$\begin{aligned} 0_i &= s_i p_i \\ &= s_i^* \, p_i^* \gamma_i \alpha_i \, , \end{aligned}$$

so,

$$p_i^* = \frac{0_i}{s_i^* \gamma_i \alpha_i}$$

$$= \left(\frac{\gamma_i a_i}{\gamma_i \alpha_i}\right)\left(\frac{0_i}{I_i - 0_i}\right) .$$

One may verify that substituting the above expressions for s_i^* and p_i^* into the balance equation reduces it to an identity. □

Lemma 2. For any point $(I_i, 0_i)$ in the flow plane for grade i the vector of retraining flows $r_i = (r_{i1}, \cdots, r_{iK})$ is well defined, and the retraining flow for group k in grade i, r_{ik}, is given by

$$r_{ik}(I_i, 0_i) = g_{ik} - I_i \left(\frac{\gamma_{ik} a_{ik}}{\gamma_i a_i}\right) - 0_i\left(\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i \alpha_i} - \frac{\gamma_{ik} a_{ik}}{\gamma_i a_i}\right) .$$

Proof. From Section G, equation (14) gives r_{ik} as a function of s_i^*,

$$r_{ik}(s_i^*) = g_i \bar{l}\left(\frac{g_{ik}}{g_i \bar{l}} - \frac{\gamma_{ik}\alpha_{ik}}{\gamma_i \alpha_i}\right) + s_i^* \gamma_i a_i \left(\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i \alpha_i} - \frac{\gamma_{ik} a_{ik}}{\gamma_i a_i}\right) .$$

From Lemma 1 we have

$$s_i^* \gamma_i a_i = I_i - 0_i \, .$$

Thus,

$$r_{ik}(I_i, 0_i) = I_i\left(\frac{g_{ik}}{I_i} - \frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i}\right) + (I_i - 0_i)\left(\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} - \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i}\right)$$

$$= g_{ik} - I_i\left(\frac{\gamma_{ik}a_{ik}}{\gamma_i a_i}\right) - 0_i\left(\frac{\gamma_{ik}\alpha_{ik}}{\gamma_i\alpha_i} - \frac{\gamma_{ik}a_{ik}}{\gamma_i a_i}\right).$$

We must establish that g_{ik} is a well defined function of I_i. Let e_k be a K x 1 vector having all zero components except for a 1 in its k^{th} component. Let

$$\tilde{g}_i = \tilde{f}_1, \quad \text{when } i=1,$$

$$= \frac{\gamma_{i-1} @ \alpha_{i-1}}{\gamma_{i-1}\alpha_{i-1}}, \quad \text{when } i=2,\cdots,n.$$

From the introductory discussion the distribution vector $\tilde{g}_i$ is known.

Then $g_i = (g_i\bar{1})\tilde{g}_i = I_i\tilde{g}_i$,

and $g_{ik} = g_i e_k$

$$= (g_i\bar{1})\tilde{g}_i e_k$$

$$= I_i\tilde{g}_i e_k \,. \qquad \square$$

Lemmas 1 and 2 demonstrate that total stock s_i^* , overall promotion rate p_i^* and the retraining flows r_i are uniquely determined by the total flows I_i and 0_i. It follows that the optimal retraining cost in grade i is also uniquely determined by these flows.

The following notation is useful in the proof of theorem 3 and in subsequent analysis:

1) $\tilde{g}_{ik} = \tilde{g}_i e_k$

where $\tilde{g}_i$ and e_k are as defined in the proof of Lemma 2,

2) $b_{ik} = \tilde{g}_{ik} - \frac{\gamma_{ik} a_{ik}}{\gamma_i a_i}$,

3) $\beta_{ik} = \frac{\gamma_{ik} a_{ik}}{\gamma_i a_i} - \frac{\gamma_{ik} \alpha_{ik}}{\gamma_i \alpha_i}$.

We note that $\tilde{g}_{ik}$, b_{ik} and β_{ik} are scalars; their values are known and they do not depend on I_i or 0_i.

We then have

$$r_{ik}(I_i, 0_i) = I_i\left(\tilde{g}_{ik} - \frac{\gamma_{ik} a_{ik}}{\gamma_i a_i}\right) + 0_i\left(\frac{\gamma_{ik} a_{ik}}{\gamma_i a_i} - \frac{\gamma_{ik} \alpha_{ik}}{\gamma_i \alpha_i}\right)$$

$$= b_{ik} I_i + \beta_{ik} 0_i . \qquad (19)$$

Note that the retraining flows in grade i are linear functions of the total flows into and out of grade i.

For any point $(I_i, 0_i)$ in the flow plane (excluding the origin) define the ray through $(I_i, 0_i)$ as $R(I_i, 0_i)$, i.e., $R(I_i, 0_i) = \{(x,y): \text{ for some } c>0,\ x=cI_i,\ y = c0_i\}$.

Theorem 3. For any point $(I_i, 0_i)$ in the flow plane for grade i (excluding the origin),

1) the dual optimal solution to the retraining transportation problem, v_i^o, is the same for all points in $R(I_i, 0_i)$,

2) the value of the optimal retraining cost for any point in $R(I_i, 0_i)$ is

$$d_i^o(cI_i, c0_i) = cd_i^o(I_i, 0_i), \ (c>0),$$

3) the value of p_i^* is constant in $R(I_i, 0_i)$.

Proof. We begin by writing the dual retraining transportation problem with the retraining flows as functions of the total flows into and out of grade i.

$$\text{DRTP]} \quad \max d_i(I_i, 0_i) = \sum_{k=1}^{K} r_{ik}(I_k, 0_k)\, v_{ik}$$

$$\text{ST} \quad v_{ik} - v_{im} \leq c_i(k,m);\ k \varepsilon K_i^+,\ m\ K_k^-$$

$$v_{ik},\ v_{im} \quad \text{UNRESTRICTED.}$$

From Lemma 2 and equation (19),

$$r_{ik}\ (I_i, 0_i) = b_{ik}\ I_i + \beta_{ik}\ 0_i\,.$$

The objective function in DRTP may then be written,

$$d_i(I_i, 0_i) = \sum_{k=1}^{K} (b_{ik} I_i + \beta_{ik}\ 0_i) v_{ik}\,.$$

Let v_i be a solution to DRTP at $(I_i, 0_i)$, i.e.,

$$d_i^o(I_i, 0_i) = \sum_{k=1}^{K} (b_{ik}\ I_i + \beta_{ik}\ 0_i) v_{ik}\,.$$

At any other point in $R(I_i, 0_i)$ the objective function in DRTP is,

$$d_i(cI_i, c0_i) = \sum_{k=1}^{k} (b_{ik}\ cI_i + \beta_{ik}\ c0_i) v_{ik}$$

$$= c \sum_{k=1}^{k} (b_{ik}\ I_i + \beta_{ik}\ 0_i) v_{ik}$$

$$= cd_i(I_i, 0_i)\,.$$

It is well-known that multiplying the objective function of a linear program by a positive constant c,

1) does not change the value of the variables in the optimal solution,

2) changes the optimal value of the objective function by the factor c.

Thus we have proven parts 1) and 2) of Theorem 3.

To prove part 3) we note that from Lemma 1 the value of p_i^* at any point in $R(I_i, O_i)$ is,

$$p_i^* = \left(\frac{\gamma_i a_i}{\gamma_i \alpha_i}\right)\left(\frac{cO_i}{cI_i - cO_i}\right), \quad \text{where } c>0,$$

$$= \left(\frac{\gamma_i a_i}{\gamma_i \alpha_i}\right)\left(\frac{O_i}{I_i - O_i}\right),$$

i.e., the value of p_i^* is constant in $R(I_i, O_i)$. $\square$

The flow plane for grade i is sketched in Figure 3. The constant-value curves for s_i^* and p_i^* are depicted by dashed and solid lines respectively. Note that the ray on which $I_i = O_i$ corresponds to $p_i^* = \infty$ and $s_i^* = 0$. The ray on which $O_i = 0$ corresponds to $p_i^* = 0$.

We see from Theorem 3 that the optimal retraining cost is linear on any ray in the flow plane. Next, we consider the character of the optimal retraining cost in the region bounded by two rays, i.e., a cone in the flow plane.

Lemma 4. If v_i^o is the dual optimal solution to the retraining transportation problem at two points in the flow plane, (I_i', O_i') and (I_i'', O_i''), then v_i^o is the dual optimal solution at any point (I_i, O_i) such that for $0 \leq \theta \leq 1$,

$$(I_i, O_i) = \theta(I_i', O_i') + (1-\theta)(I_i'', O_i'') .$$

Proof. If v_i^o solves the dual retraining transportation problem (DRTP) at (I_i', O_i') and (I_i', O_i'), then,

1) v_i^o is a feasible solution to DRTP everywhere in the flow plane for grade i,

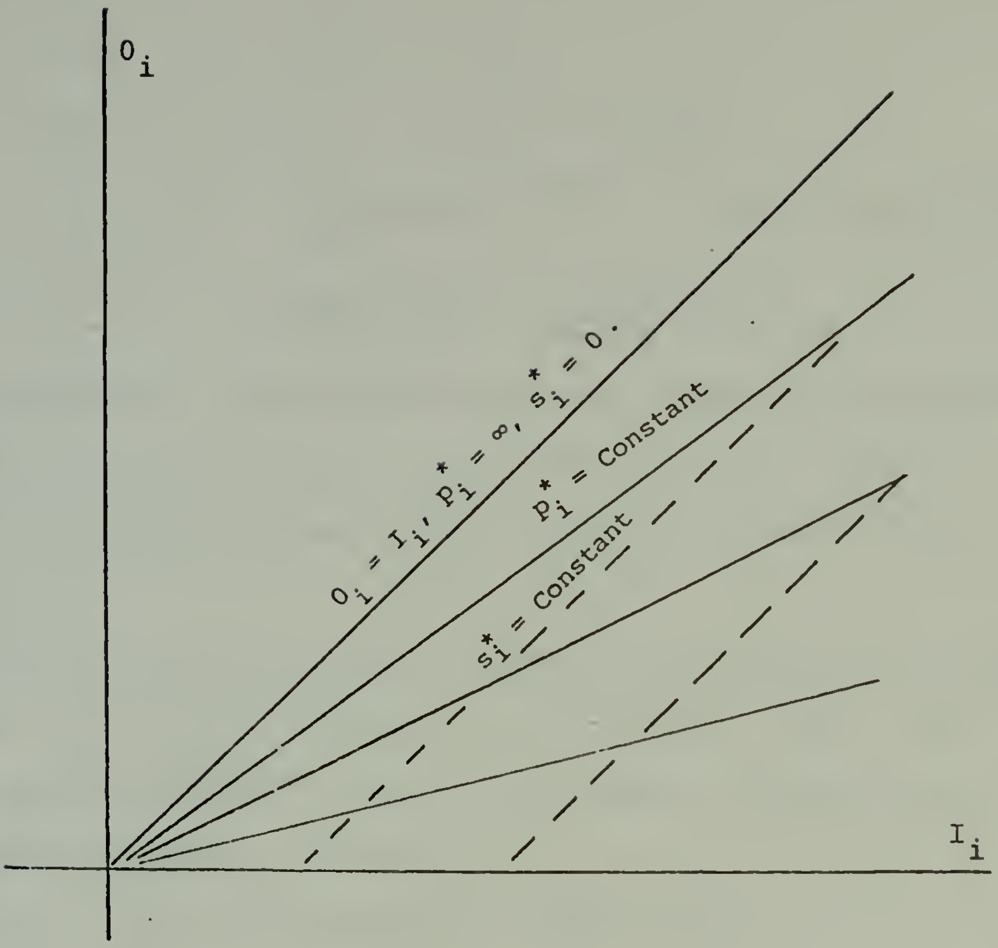

Figure 3. Constant-Value Curves for s_i^* and p_i^* in the Flow Plane for Grade i.

$$2)\quad d_i^o(I_i', O_i') = \sum_{k=1}^{K} (b_{ik} I_i' + \beta_{ik}\ O_i') v_{ik}^o$$

$$3)\quad d^o(I_i'',\ O_i'') = \sum_{k=1}^{K} (b_{ik}\ I_i'' + \beta_{ik}\ O_i'')\ v_{ik}^o .$$

Let,

$$d_i(I_i,\ O_i;\ v_i^o) = \sum_{k=1}^{K} (b_{ik} I_i + \beta_{ik}\ O_i) v_{ik}^o . \tag{20}$$

We must show for $0 \leq \theta \leq 1$, and

$$(I_i,\ 0_i) = \theta(I_i',\ 0_i') + (1-\theta)(I_i'',\ 0_i''), \tag{21}$$

that $d_i^o(I_i,\ 0_i) = d_i(I_i,\ 0_i;\ v_i^o)$.

By substituting (21) into (20) we find,

$$d_i(I_i,\ 0_i;\ v_i^o) = \sum_{k=1}^{K} (b_{ik}[\theta I_i' + (1-\theta)I_i''] + \beta_{ik}[\theta 0_i + (1-\theta)0_i''])v_{ik}^o$$

$$= \theta d_i^o(I_i',\ 0_i') + (1-\theta)d_i^o(I_i,\ 0_i'').$$

Because the dual optimal solution implies a basis for the primal optimal solution, it is readily verified that the sets K_i^+ and K_i^- cannot change unless the dual optimal solution changes.

Let x_i' solve the primal retraining transportation problem (PRTP) at $(I_i',\ 0_i')$, and let x_i'' solve PRTP at $(I_i'',\ 0_i'')$. Then because the optimal values of the primal and dual linear programs are equal, we have,

$$\sum_{k\varepsilon K_i^-} \sum_{m\varepsilon K_i^-} c_i(k,m)x_i'(k,m) = d_i^o(I_i',\ 0_i'), \tag{22}$$

$$\sum_{k\varepsilon K_i^-} \sum_{m\varepsilon K_i^-} c_i(k,m)x_i''(k,m) = d_i^o(I_i'',\ 0_i''). \tag{23}$$

$$x_i(k,m) = \theta x_i'(k,m) + (1-\theta)x_i''(k,m),\ k\varepsilon K_i^+,\ m\varepsilon K_i\ .$$

By hypothesis $\{x_i'(k,m)\}$ and $\{x_i''(k,m)\}$ are feasible in PRTP. Thus

$$\sum_{m\varepsilon K^-} x_i(k,m) = \sum_{m\varepsilon K_i^-} (\theta x_i'(k,m) + (1-\theta)x_i''(k,m)), \quad k\theta K_i^+$$

$$= \theta\sum_{m\varepsilon K_i^-} x_i'(k,m) + (1-\theta)\sum_{m\varepsilon K_i^-} x_i''(k,m)$$

$$= \theta r_{ik}(I_i', 0_i') + (1-\theta)r_{ik}(I_i'', 0_i'').$$

Using equations (19) and (21) we have,

$$\sum_{m\varepsilon K_i^-} x_i(k,m) = r_{ik}(I_i, 0_i), \quad k\varepsilon K_i^+ .$$

Similarly it can be verified that

$$\sum_{k\varepsilon K_i^+} x_i(k,m) = -r_{ik}(I_i, 0_i), \quad m\varepsilon K_i^- .$$

Furthermore, $x_i'(k,m) \geq 0$, $x_i''(k,m) \geq 0$ and $0 \leq \theta \leq 1$ implies that

$$\theta x_i'(k,m) + (1-\theta)x_i''(k,m) \geq 0 ,$$

i.e., $x_i(k,m) \geq 0$. We note that this is the only part of the proof that uses the condition $0 \leq \theta \leq 1$.

Thus we have shown that $x_i(k,m)$ is feasible in PRTP. Using (20), (22) and (23) we find that

$$\sum_{k\varepsilon K_i^+} \sum_{m\varepsilon K_i^-} c_i(k,m)x_i(k,m) = \theta d_i^o(I_i', 0_i') + (1-\theta)d_i^o(I_i'', 0_i'')$$

$$= d_i(I_i, 0_i, v_i^o).$$

So we have found a primal feasible solution $\{x_i(k,m)\}$ and a dual feasible solution v_i^o and both give the same value of the objective function, so they are optimal solutions. In particular, v_i^o is the dual optimal solution at $(I_i, 0_i)$ as was to be shown. □

If solution of the retraining transportation at two points in the flow plane results in the same dual optimal solution v_i^o, then the proof of Lemma 4 enables us to determine the extent of the cone in the flow plane in which the dual optimal solution does not change.

Corollary 5. If two points in the flow plane, $(I_i', 0_i')$ and $(I_i'', 0_i'')$, have solutions to PRTP, x_i' and x_i'' respectively, and have the same solution to DRTP, v_i^o, then for any θ such that

$$\theta x_i' + (1-\theta)x_i'' \geq \bar{0},$$

the solutions to PRTP and DRTP at

$$(I_i, 0_i) = \theta(I_i', 0_i') + (1-\theta)(I_i'', 0_i'')$$

are,

$$\theta x_i' + (1-\theta)x_i'',$$

and,

$$v_i^o$$

respectively.

Proof. Follows from the proof of Lemma 5. ☐

The foregoing results enable us to describe the optimal retraining costs as a function of the total flows into and out of grade i.

Theorem 6. The optimal retraining cost in grade i is a piecewise linear convex function of the total flows into and out of grade i.

Proof. The piecewise linearity of the optimal retraining costs follows from equation (19), Lemma 4 and Corollary 5.

To show the convexity of the optimal costs, it will be convenient to use the original retraining cost minimization program developed in Section F.

$$\text{P1]}\quad \min \sum_{k=1}^{K} \sum_{m=1}^{K} c_i(k,m)x_i(k,m)$$

$$\text{ST} \quad \sum_{m=1}^{K} (x_i(k,m) - x_i(m,k)) = r_{ik},\ k=1,\cdots,K$$

$$x_i(k,m) \geq 0 \ .$$

Suppose x_i' solves P1 at $(I_i', 0_i')$, and x_i'' solves P1 at $(I_i'', 0_i'')$. Then,

$$d_i^o(I_i' \quad 0_i') = \sum_{k=1}^{K} \sum_{m=1}^{K} c_i(k,m)x_i'(k,m), \tag{25}$$

and

$$d_i^o(I_i'', 0_i'') = \sum_{k=1}^{K} \sum_{m=1}^{K} c_i(k,m)x_i''(k,m). \tag{26}$$

Let,

$$x_i = \theta x_i' + (1-\theta)x_i''$$

Then $x_i \geq \bar{0}$, and,

$$\begin{aligned}
\sum_{m=1}^{K} (x_i(k,m)-x_i(m,k)) &= \sum_{m=1}^{K} [\theta(x_i'(k,m)-x_i'(m,k) + (1-\theta)(x_i''(k,m) - x_i''(m,k))] \\
&= \theta \sum_{m=1}^{K} (x_i'(k,m)-x_i'(m,k)) + (1-\theta) \sum_{m=1}^{K} (x_i''(k,m)-x_i''(m,k)) \\
&= \theta r_{ik}(I_i', 0_i') + (1-\theta)r_{ik}(I_i'', 0_i'') \\
&= r_{ik}(I_i, 0_i).
\end{aligned}$$

So x_i is a feasible solution to P1 at $(I_i, 0_i)$. The value of the objective function with this solution is,

$$\sum_{k=1}^{K} \sum_{m=1}^{K} c_i(k,m)x_i(k,m) = \sum_{k=1}^{K} \sum_{m=1}^{K} c_i(k,m)[\theta x_i'(k,m) + (1-\theta)x_i''(k,m)]$$

$$= \theta d_i^o(I_i', 0_i') + (1-\theta)d_i^o(I_i'', 0_i'')$$

from (25) and (26).

But the minimal value of the objective function can be no greater than the value resulting from the feasible solution x_i, i.e.,

$$d_i^o(I_i, 0_i) \leq \theta d_i^o(I_i', 0_i') + (1-\theta)d_i^o(I_i'', 0_i''). \qquad \square$$

We use the assumption that $0_i = I_{i-1}$, $i=1,\cdots,n-1$ to combine the flow planes.

Let us define an (n+1)-dimensional "flow space" as the nonnegative orthant of E^{n+1}, and let the i^{th} component of any point in this space be the total flow into grade i, $i=1,\cdots,n$, and the $(n+1)^{st}$ component is identically zero. Let,

$$d^o(I_1,\cdots,I_{n+1}) = \sum_{i=1}^{n} d_i^o(I_i, I_{i+1}), \qquad [27]$$

where $I_{n+1} \equiv 0$.

Theorem 7. The function $d^o(I_1,\cdots,I_{n+1})$ is convex. It is linear on cones in which the optimal dual solutions do not change.

Proof. The sum of convex functions is a convex function, so Theorem 6 and the definition of $d^o(I_1,\cdots,I_{n+1})$ indicate it is convex.

We have from Corollary 6 that in each grade i, the optimal retraining cost $d_i^o(I_i, I_{i+1})$ is linear in any region in which the dual optimal solution v_i^o does not change.

Theorem 3 and Lemma 4 indicate that such regions are two-dimensional cones. The intersection of such two-dimensional cones in the (n+1)-dimensional flow space defines a cone in E^{n+1}. The sum of linear functions is linear, so the latter part of the theorem follows. ☐

It has been shown in this section that under the assumption that external flows are restricted to the lowest grade, the retraining costs are a convex function of the total flows of people between grades. Consequently, in a cost minimizing scheme such as that described in Section H, any local optimum is also a global optimum provided the set of solutions considered feasible is a convex set in the flow space.

APPENDIX A. SUMMARY OF NOTATION

For the most part scalars and vectors are denoted by lower case letters, and matrices are denoted by upper case letters. There are exceptions.

The subscript i is used to denote the first characteristic, grade. The subscript j is used to denote the second characteristic.

A matrix of zeroes is denoted by $\bar{\bar{0}}$. A vector of zeroes is denoted by $\bar{0}$. A vector of ones is denoted by $\bar{1}$. The dimensions of $\bar{\bar{0}}$, $\bar{0}$ and $\bar{1}$ are implied by the context in which they are used. For any vector x we use $x\bar{1}$ to denote the sum of the components of x.

Unless otherwise indicated the empty sum and the empty product are defined as identity elements for the corresponding operator. For example, if B_m is a matrix then,

$$\sum_{m=k}^{i-1} B_m = \bar{\bar{0}}, \text{ the additive identity matrix, when } k \geq 1,$$

$$\prod_{m=k}^{i-1} B_m = I, \text{ the multiplicative identity matrix, when } k \geq i.$$

When a matrix is displayed in terms of its submatrices, the zero submatrices may not be shown explicitly. For example,

$$Q = \begin{bmatrix} Q_1 & P_1 & \\ & Q_2 & P_2 \\ & & Q_3 \end{bmatrix} = \begin{bmatrix} Q_1 & P_1 & \bar{\bar{0}} \\ \bar{\bar{0}} & Q_2 & P_2 \\ \bar{\bar{0}} & \bar{\bar{0}} & Q_3 \end{bmatrix}$$

For any square matrix A we use A_{dg} to denote the matrix which has the same elements as A on its main diagonal but has zeroes elsewhere. For any matrix B we use B_{sq} to denote the matrix which has the square of the elements of B as its elements.

The matrix P_i and the vector $s_i(t)$ are defined for $i=1,\cdots,n$. For convenience we may use the convention $P_i=\bar{\bar{0}}$ and $s_i(t) = \bar{0}$ for $i < 1$.

If x and y are two vectors such that

$$x = (x_1,\cdots,x_m)$$

and,

$$y = (y_1,\cdots,y_m),$$

then we define the congruent multiplication operator, @, by,

$$x@y = (x_1y_1,\cdots,x_my_m).$$

Congruent multiplication is discussed in some detail in R. A. Howard, 1970.

A "distribution vector" is a nonnegative vector having components that sum to one. Row vectors are enclosed in parentheses. Column vectors are enclosed in brackets.

The symbol $\square$ denotes the end of a proof.

APPENDIX B. DIAGONAL MATRICES

Let A be a matrix, not necessarily square, and let A (k,j) be the element of A in row i and column j. We say that A is a diagonal matrix if for some integer k_A, A $(k,j) \neq 0$ implies that $j-i$ equals k_A. We call k_A the index of the diagonal matrix A.

The key cell of the diagonal matrix A is defined as:

(1) A $(1,1)$ if $k_A = 0$,

(2) A $(1,k_A + 1)$ if $k_A > 0$,

(3) A $(1-k_A, 1)$ if $k_A < 0$.

We will use the notation (i_A, j_A) for the row and column of the key cell of A. Note that $j_A - i_A = k_A$.

If A is a zero matrix, then it is a diagonal matrix, and the value of its index is arbitrary. If A is a non-zero diagonal matrix, then its index is unique and its key cell is well defined.

Let A be a non-zero m x n diagonal matrix, and let Z_A be maximum number of non-zero elements in A. Then it is readily verified that

$$z_A = \min \{m + \min \{0, K_A\}, \quad n - \max \{0, k_A\}\}.$$

All non-zero elements of A are in the set:

$$\{A(i_A + r, j_A + r); \quad r = 0,1,\cdots,z_A - 1\}.$$

This set is called the non-zero diagonal of A.

The purpose of this appendix is to show that diagonal matrices can be stored and multiplied in compacted form.

If A is an m x n diagonal matrix, then the compacted form of A is a vector A^* where,

$A^*(1) = m,$

$A^*(2) = n,$

$A^*(3) = k_A,$

$A^*(4) = z_A,$

and the fifth through last elements of A^* are the elements of the non-zero diagonal of A.

A zero matrix A may be stored as a four-component vector A^*:

$A^*(1) = m$

$A^*(2) = n$

$A^*(3) \geq n$ or $A^*(3) \leq -m$

$A^*(4) \leq 0.$

We will show that the product of diagonal matrices is a diagonal matrix, and that the product is efficiently computed using the compacted forms of the multiplicand and product matrices.

Lemma 1. If A and B are diagonal matrices such that C = AB is defined, then C is a diagonal matrix. Furthermore the index of C is the sum of the indices of A and B.

Proof. If C is a zero matrix, then it is a diagonal matrix and the value of its index is arbitrary.

If C is not a zero matrix, then for some row i and column j,

$C(i, j) \neq 0.$

We note that $C(i, j)$ is the inner product of the i^{th} row of A and the j^{th} column of B, and that each of the rows and columns of A and B contain at most one non-zero element. Consequently, there must exist nonnegative integers r_A and r_B such that:

$$C(i, j) = A(i_A + r_A, j_A = r_A)\ B(i_B + r_B, j_B + r_B)$$

$$i = i_A + r_A$$

$$j = j_B + r_B$$

$$j_A = r_A = i_B + r_B .$$

Starting with the last equation

$$
\begin{aligned}
r_B &= j_A + r_A - i_B \\
j &= j_B + j_A + r_A - i_B \\
j-i &= j_B + j_A + i_B - i_A \\
&= (j_A - i_A) + (j_B - i_B) \\
&= k_A + k_B.
\end{aligned}
$$

$\square$

Continuing with notation used above, we see that each element of the non-zero diagonal of C is the product of elements of the non-zero diagonals of A and B. We consider next how the elements of the non-zero diagonals of A and B are paired to form the non-zero diagonal of C.

From the proof of the lemma, if $C\ (i,j) \neq 0$, then for some non-negative integers r_A and r_B:

(1) $C(i,j) = A(i_A + r_A, j_A + r_A)\ B(i_B + r_B, j_B + r_B)$

(2) $i = i_A + r_A$

(3) $j = j_B + r_B$

(4) $j_A + r_A = i_B + r_B$

To find the relation between r_A and r_B we note that

$$\begin{aligned}(j_B + r_B) - (i_A + r_A) &= j - i \\ &= k_A + k_B,\end{aligned}$$

so,

$$\begin{aligned}r_B - r_A &= k_A + k_B - j_B + i_A \\ &= (k_B - j_B) + (i_A + k_A)\end{aligned}$$

But,

$$\begin{aligned}i_A + k_A &= 1 + k_A & \text{if } k_A > 0 \\ &= 1 & \text{if } k_A \leq 0,\end{aligned}$$

and,

$$\begin{aligned}k_B - j_B &= -1 & \text{if } k_B \geq 0 \\ &= k_B - 1 & \text{if } k_B < 0.\end{aligned}$$

Thus,

$$r_B - r_A = k_B^- + k_A^+$$

$$\begin{aligned}\text{where } k_A^+ &= k_A \text{ if } k_A > 0 \\ &= 0 \quad \text{otherwise,} \\ k_B^- &= k_B \text{ if } k_B < 0 \\ &= 0 \text{ otherwise.}\end{aligned}$$

Let $d_{AB} = k_B^- + k_A^+$.

1) If $k_B^- + k_A^+ = d_{AB} \geq 0$, then $r_B = r_A + d_{AB} \geq 0$, and the first element of the non-zero diagonal of A is paired with the $(d_{AB} + 1)^{st}$ element of the non-zero diagonal of B.

2) If $k_B^- + k_A^+ = d_{AB} < 0$, then $r_B = r_B - d_{AB} > 0$, and the first element of the non-zero diagonal of B is paired with the $(1-d_{AB})^{th}$ element of the non-zero diagonal of A.

Consider the case

$$r_A = -d_{AB}^-$$

$$r_B = d_{AB}^+$$

Then,

$$r_B - r_A = d_{AB}^+ + d_{AB}^- = d_{AB} ,$$

and,

$$C(i,j) = A(i_A - d_{AB}^-, j_A - d_{AB}^-)B(i_B + d_{AB}^+, j_B + d_{AB}^+)$$

$$i = i_A - d_{AB}^-$$

$$j = j_B + d_{AB}^+ .$$

If $k_A \geq 0$ or $k_B \leq 0$, then,

1) $d_{AB} < 0 \rightarrow k_B < 0$

$\rightarrow j_B = 1$

$\rightarrow j = 1$

2) $d_{AB} > 0 \rightarrow k_A > 0$

$\rightarrow i_A = 1$

$\rightarrow i = 1$

3) $d_{AB} = 0 \rightarrow k_A = k_B = 0$

$\rightarrow i_A = j_B = 1$

$\rightarrow i = j = 1$

So if $k_A \geq 0$ or $k_B \leq 0$, then either i or j equals 1 in the equation

$$C(i,j) = A(i_A - d_{AB}^-, j_A - d_{AB}^-)B(i_B + d_{AB}^+, j_B + d_{AB}^+) .$$

That is, C(i,j) is the key element of C.

If $K_A < 0$ and $k_B > 0$, then $d_{AB} = 0$ and

$C(i,j) = C(i_A, j_B) = C(1-k_A, 1+k_B)$,

and the first min $\{-k_A, k_B\}$ elements of the non-zero diagonal of C are zeroes.

The foregoing observations lead to the following algorithm for multiplying two diagonal matrices.

1) Compute $d_{AB} = k_B^- + k_A^+$.

2) Drop the first d_{AB}^- elements of the non-zero diagonal of A and store the remaining elements in a vector VA.

3) Drop the first d_{AB}^+ elements of the non-zero diagonal of B and store the remaining elements in a vector VB.

4) Drop elements from the end of the longer of the two vectors VA and VB until they are of equal length. Multiply corresponding components of these vectors to form a new vector VC.

5) Append max $\{0, \min \{-k_A, k_B\}\}$ zeroes in front of VC. Compute z_C, the length of the non-zero diagonal of C, and append zeroes to the end of VC as necessary to form the non-zero diagonal of C.

6) The first four elements of the compacted form of C are A(1), B(2), $k_A + k_B$, z_C. The remaining elements are the non-zero diagonal of C as computed in 5).

LIST OF REFERENCES

Bartholomew, D. J., "Note on the Measurement and Prediction of Labour Turnover," J. R. Stat. Soc., Vol. A122, pp. 232-239, 1959.

Bartholomew, D. J., Stochastic Models for Social Processes, Wiley, 1967.

Blumen, I., Hogan, M., and McCarthy, P., The Industrial Mobility of Labor as a Probability Process. Cornell University, 1955.

Burke, C. K. and Rosenblatt, M., "A Markovian Function of a Markov Chain," Annals of Math. Stat., Vol. 29, pp. 1112-1122, 1958.

Charnes, A., Cooper, W. W., and Niehaus, R. J., Studies in Manpower Planning, Office of Civilian Manpower Management, Department of the Navy, 1972.

Dantzig, G. B., Linear Programming and Extensions, pp. 270-271, Princeton University Press, 1963.

Debreu, G., and Herstein, I. N., "Nonnegative Square Matrices," Econometrica, Vol. 21, No. 4, pp. 597-607, 1953.

Grinold, R. C. and Marshall, K. T., Manpower Planning Models, to be published.

Grinold, R. C., Marshall, K. T. and Oliver, R. M., "Longitudinal Manpower Planning Models," Research Report ORC73-15, Operations Research Center, University of California, Berkeley, 1973.

Grinold, R. C. and Stanford, R. E., "Optimal Control of a Graded Manpower System," Man. Sci., Vol. 20, No. 8, pp. 1201-1216, 1974.

Halmos, P. R., Lectures on Ergodic Theory, Chelsea, 1956.

Howard, R. A., Dynamic Probabilistic Systems, Vol. I: Markov Models, Wiley, 1970.

Howard, R. A., Dynamic Probabilistic Systems, Vol. II: Semi-Markov and Decision Processes, Wiley, 1971.

Karlin, S., A First Course in Stochastic Processes, Academic Press, 1968.

Kemeny, J. G. and Snell, J. L., Finite Markov Chains, Van Nostrand, 1960.

Lane, K. F. and Andrew, J. E., "A Method of Labour Turnover Analysis," J. R. Stat. Soc., Vol. A118, pp. 296-323, 1955.

Marshall, K. T., "A Comparison of Two Personnel Prediction Models," Operations Research, Vol. 21, No. 3, pp. 810-822, 1973.

Marshall, K. T. and Oliver, R. M., "A Constant-Work Model for Student Attendance and Enrollment," Operations Research, Vol. 18, No. 2, pp. 193-206, 1970.

RAND Corporation, "Planning in Large Personnel Systems: A Reexamination of the TOPLINE Static Planning Model," R-1274-PR, 1973.

Rowland, K. M. and Sovereign, M. G., "Markov-Chain Analysis of Internal Manpower Supply," Industrial Relations, Vol. 9, No. 1, pp. 88-99, 1969.

Silcock, H., "The Phenomenon of Labour Turnover," J. R. Stat. Soc., Vol. A117, pp. 429-440, 1954.

U.S. Navy, "Computer Models for Manpower and Personnel Management: State of Current Technology," (NAMPS Project Report 73-2), Naval Personnel Research and Development Laboratory, 1973.

Young, A. and Almond, G., "Predicting Distributions of Staff," Comp. J., Vol. 3, pp. 246-250, 1961.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Department of Operations Research and Administrative Sciences, Code 55 Naval Postgraduate School Monterey, California 93940	1
4.	Chief of Naval Personnel Pers 11b Department of the Navy Washington, D.C. 20370	1
5.	Assoc. Prof. G. T. Howard, Code 55 Hk Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
6.	Assoc. Prof. K. T. Marshall, Code 55 Mt Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	5
7.	Assoc. Prof. P. R. Milch, Code 55 Mh Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
8.	Assoc. Prof. T. Jayachandran, Code 53 Jy Department of Mathematics Naval Postgraduate School Monterey, California 93940	1
9.	Assoc. Prof. G. A. Myers, Code 52 Mv Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1

10. Assoc. Prof. Gordon Bradley, Code 55 Bz 1
Department of Operations Research
and Administrative Sciences
Naval Postgraduate School
Monterey, California 93940

11. Assoc. Prof. R. W. Butterworth, Code 55 Bd 1
Department of Operations Research
and Administrative Sciences
Naval Postgraduate School
Monterey, California 93940

12. LCDR William J. Hayne, USN 4
350 Rose Lane
St. Paul, Minnesota 55117

13. Dr. Thomas Varley, Code 434 1
Program Director, Operations Research
Office of Naval Research
Arlington, Virginia 22217

14. Dr. Robert M. Oliver 1
Director, Operations Research Center
Etcheverry Hall
University of California
Berkeley, California 94720

15. Dr. Richard C. Grinold 1
C.O.R.E.
De Croylaan 54
3030 Heverlee
Belgium

16. Commandant, Marine Corps 1
Code MPP
Headquarters Marine Corps
Washington, D.C. 20380

17. Dr. D. J. Bartholomew 1
London School of Economics
London, WC 1
England

Thesis
H4064 Hayne 152670
c.1
Analysis of some two-characteristic Markov-type manpower flow models with retraining applications.

2 MAY 77
3 DEC 78
11 APR 79
18 JUN 81
2 JUN 82
JUN 24 85
28 OCT 86

24125
25028
26[illegible]
26405
27730
33093
31167

Thesis
H4064 Hayne 152670
c.1
Analysis of some two-characteristic Markov-type manpower flow models with retraining applications.

thesH4064
Analysis of some two-characteristic Mark
3 2768 001 02069 6
DUDLEY KNOX LIBRARY